

ajer

THE ALBERTA JOURNAL OF
EDUCATIONAL RESEARCH

VOLUME XXV

NUMBER 2

JUNE 1979

PUBLISHED BY
THE UNIVERSITY OF ALBERTA • EDMONTON

THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

A quarterly journal devoted to the dissemination, criticism, interpretation and encouragement of all forms of systematic enquiry into education and fields related to or associated with education.

Published quarterly in March, June, September, December by the
Faculty of Education, The University of Alberta

CONSULTING EDITORS

J. Britton
University of London

J. Calam
The University of British Columbia

M. Connelly
*The Ontario Institute for
Studies in Education*

K. De Clerck
State University of Ghent

R. N. Evans
*University of Illinois at
Urbana-Champaign*

R. H. Farquhar
University of Saskatchewan

E. Gagné
University of Ottawa

G. Harman
University of Melbourne, Australia

S. Hunka
The University of Alberta

J. W. G. Ivany
Simon Fraser University

D. A. MacIver
University of New Brunswick

L. D. Nelson
The University of Alberta

W. C. Nesbit
Memorial University of Newfoundland

E. Pedersen
McGill University

EDITOR: H. W. Hodysh

SECRETARY: A. Onishenko

FACULTY PUBLICATIONS COMMITTEE

M. A. Assheton-Smith
T. P. Atkinson
N. C. Bhattacharya
W. T. Fagan

H. W. Hodysh
W. W. Laing
R. G. Martin

E. Miklos (Chairman)
J. W. Osborne
C. H. Preitz

Editorial policy and the discussion and disposition of manuscripts are the joint responsibility of the Publications Committee. The views expressed and the accuracy of the statements made are the responsibility of the individual authors. The editor is solely responsible for the editorial comments.

AJER gratefully acknowledges support from the Social Sciences and Humanities Research Council of Canada and from the Alberta Advisory Committee for Educational Studies.

AJER is indexed in the *Canadian Education Index*, *Current Contents/Social and Behavioral Sciences*, and *Social Science Citation Index*; appropriate articles are abstracted in *Educational Administration Abstracts*, *Psychological Abstracts*, *Sociology of Education Abstracts*, and *Language Behavior Abstracts*.

The subscription rate is \$8.00 per year; single copies are \$2.50 each. Please make cheques payable to *The Alberta Journal of Educational Research*. All back issues are available; rates supplied on request. Claims for undelivered copies must be received within three months of the month of publication.

Address all communications and manuscript submissions to the Editor, *The Alberta Journal of Educational Research*, Faculty of Education, 732 Education South, The University of Alberta, Edmonton, Canada, T6G 2G5.

SECOND CLASS MAIL REGISTRATION NUMBER 1436



The Alberta Journal of Educational Research

Volume XXV, Number 2

June, 1979

CONTENTS

Achievement and the Prediction of Achievement in English First and Second Language Children	61
<i>B. Panunto and D. White</i>	
Facilitating Reading Comprehension Through Text Structure Manipulation	68
<i>J. M. Mason and J. R. Kendall</i>	
School-Centered Community Conflict: The Holdeman Mennonite Case in Alberta	77
<i>M. Assheton-Smith and K. Toohey</i>	
A Study of Moral Education in Surrey, B.C. Secondary Schools	89
<i>D. M. Williams</i>	
The Effects of Attentional Focus on Graphic Discrimination	103
<i>E. Shaw</i>	
Using Qualitative Data in Formative Evaluation	117
<i>G. S. Aikenhead</i>	

BOOK REVIEWS

<i>Research Into Teaching Methods in Higher Education</i> , Ruth M. Beard, Donald A. Bligh and Alan G. Harding	130
Reviewed by A. MacKay	
<i>Gladly Would He Teach</i> , John W. Chalmers	131
Reviewed by W. J. Bennett	
<i>The Process of Thinking</i> , Marc Belth	132
Reviewed by F. N. Walker	

LIST OF BOOKS RECEIVED	135
------------------------------	-----

FACULTY OF EDUCATION
The University of Alberta

BRENDA PANUNTO

and

DONNA WHITE

Concordia University

Achievement and the Prediction of Achievement in English First and Second Language Children

The purpose of this study was to examine the achievement of Italian first language children in grade one and the predictive relationship between measures of ability and achievement. Subjects were 29 monolingual English-speaking children and 59 English second language children attending grade one classes in middle class areas. The Peabody Picture Vocabulary Test and the Raven Matrices had been administered to all subjects while in kindergarten. In May of the first-grade year, the subjects were given the Metropolitan Achievement Test. No significant differences were found between English first and second language groups on the achievement subtests: the second language children obtained average scores. For Italian children, the best predictor of achievement was the Peabody test given in November of the kindergarten year. (Mrs. Panunto is a Consultant to the Montreal Catholic School Commission; Dr. White is Associate Professor of Education and Psychology at Concordia University.)

Very little research is available on the school achievement of children for whom English is a second language. The studies which do exist (Ashby, Morrison & Butcher, 1970; Bhatnagar, 1970; Cohen, 1970; Little, Mabey & Whitaker, 1967; McFie & Thompson, 1970; Toronto Board of Education, 1965) indicate that such children achieve at below average levels, tend to drop out of school early, and have social and emotional adjustment problems. These results cannot be accepted without qualification for several reasons. First, these studies utilized children of many different cultural and language backgrounds. Cultural assimilation and acceptance in the school system also varied. English second language status was often confounded with low socioeconomic level. However, even under optimal conditions, it would seem reasonable to hypothesize that children for whom English is a second language would not achieve at average levels as early as grade one simply because they lack English verbal skills. One purpose of the present study was to test this hypothesis utilizing a group of Italian first language subjects

whose English vocabulary tested at the beginning and end of kindergarten was below average (White & Panunto, 1978).

The second purpose of this study was to investigate ability measures as predictors of achievement in English second language children. Achievement in school is dependent on a multitude of factors including verbal and nonverbal abilities. Research has shown that for monolingual English-speaking children, positive correlations between measures of ability and later school achievement hold over a wide range of age, intellectual, and socioeconomic levels (Dokecki, Frede & Goutney, 1969; Dudek, Goldbert, Lester & Harris, 1969; Henderson, Butler & Goffeney, 1969; Hirshoren, 1969; Lesser, Schoeninger & Bridges, 1970; Meyers, Attwell & Orpet, 1968; Mussen, Dean & Rosenbert, 1952; Panther, 1967; Plant & Southern, 1968). Both the Peabody Picture Vocabulary Test (Dunn, 1965) and the Raven Coloured Progressive Matrices (Raven, 1962, 1965) utilized in the present study have been found to correlate with later school achievement (Lesser et al, 1970; Meyers et al, 1968; Panther, 1967; Plant & Southern, 1968). There is some conflict in the literature over whether verbal or nonverbal skills are more effective predictors of achievement in monolingual subjects. Similar significant correlations have been found between both verbal and performance measures and achievement scores. There is some indication that composite measures yield the most powerful predictions for this group (Dokecki et al., 1969; Dudek et al., 1969).

It is difficult to speculate about the relative effectiveness of English verbal skills and nonverbal skills as predictors of achievement in English second language children. There is some indication that both verbal and nonverbal measures predict achievement in English second language children (Cooper, 1958; Phillipus, 1968; Rattan & MacArthur, 1968). However, no studies utilizing kindergarten-grade one subjects could be found.

Method

Subjects

The subjects were 97 children drawn from 11 schools in the English sector of the Montreal Catholic School Commission in the fall of 1975. In order to control for socioeconomic level, the schools were selected from middle class areas, and inner city schools were eliminated. From the kindergartens of the 11 schools, all available (33) monolingual English-speaking children were included in the control group. A random sample of 64 English second language children was selected for the experimental group. The experimental subjects were all of Italian origin. Children included in this group were those whose parents spoke Italian and little or no English in formal and informal contacts with the schools, and children who, according to their teachers, spoke Italian spontaneously among themselves. Given these criteria, it was inferred that Italian was the language spoken at home and that the experimental children spoke and understood Italian.

All teachers and principals described the Italian families as culturally assimilated in the sense that the parents wanted their children to learn English and to do well in school. In addition, the population of the selected schools was 75-90% Italian, thus the discomfort from cultural differences was minimized.

The mean age of both groups of children in May of 1976 was 6.1 years.

Procedure

The Peabody Picture Vocabulary Test (Dunn, 1965) was administered in November and May of each child's kindergarten year. Alternative forms of the test were used and the order of administration was counterbalanced. The English second language children performed at below average levels in both November and May, while control subjects performed at an average level. The experimental groups scored significantly lower than the control groups on both Peabody tests (White & Panunto, 1978).

The Raven Coloured Progressive Matrices (Raven, 1962, 1965) was administered individually in May of the kindergarten year. Both experimental and control subjects performed at similar, average levels on the Raven (White & Panunto, 1978).

In May of the grade one year, the Metropolitan Achievement Test, Primary 1 Battery — Form F (Durost, Bixler, Wrightstone, Prescott & Balow, 1971) was administered in groups to 88 of the original 97 subjects still in the school system.

Results

Achievement Tests

The mean scores obtained by the control and experimental groups on the subtests of the Metropolitan are shown in Table 1. It can be seen that control subjects earned higher scores on all subtests although, with the exception of mathematics, the differences seem to be slight. It should be noted that all of these mean scores represent average performance on the Metropolitan Achievement Test.

TABLE 1
ACHIEVEMENT MEANS AND *T* TESTS FOR ENGLISH FIRST
AND SECOND LANGUAGE GROUPS

Achievement Subtests	Groups		<i>t</i> test Between Groups df = 86
	English First Language	English Second Language	
Word Knowledge	23.4	20.8	1.54
Word Analysis	28.9	27.4	.94
Reading	21.3	19.4	1.02
Total Reading	45.1	40.2	1.46
Total Mathematics	50.8	37.6	1.37

In order to determine whether the differences between the mean scores obtained by the control and experimental groups were significant, a series of *t* tests were performed. As indicated in Table 1, there were no significant differences between the groups on any of the subtests. Contrary to our prediction, the Italian children in this study performed at average levels on achievement tests in grade one and were not significantly different from monolingual English controls.

Prediction of Achievement

Correlations. It was hypothesized that both verbal and nonverbal scores would correlate with achievement in both English first and second language subjects. In

TABLE 2
CORRELATION OF ABILITY AND ACHIEVEMENT MEASURES IN
ENGLISH FIRST AND SECOND LANGUAGE GROUPS

Achievement Subtests	Ability Tests		
	Peabody- Nov.	Peabody- May	Raven
<u>English First Language Group</u>			
Word Knowledge	.071	.315	.386*
Word Analysis	-.132	.063	.286
Reading	.253	.344	.483*
Total Reading	.165	.344	.429*
Total Mathematics	.278	.184	.307
<u>English Second Language Group</u>			
Word Knowledge	.568**	.400**	.264*
Word Analysis	.403**	.338**	.224
Reading	.401**	.302*	.054
Total Reading	.521**	.377**	.169
Total Mathematics	.542**	.471**	.377**

* $p < .05$

** $p < .01$

order to test this hypothesis, a Pearson r was calculated between each of the three ability scores and each of the five subtest scores of the Metropolitan. As can be seen from Table 2, the Peabody scores are significant predictors of school achievement for the Italian children tested, but are not significantly correlated with achievement measures for control subjects. The Raven scores are correlated significantly with some of the achievement subtests for both groups.

Step-wise regression. Since the independent variables have been shown to be intercorrelated (White & Panunto, 1978), a step-wise regression analysis (Nie, Bent, & Hull, 1970) for the English second language group was performed to determine whether prediction could be improved by using several independent variables. The difference between the Peabody Vocabulary scores in November and May was included as a fourth independent variable in this analysis. As can be seen in Table 3, no significant gain in predictive validity was made for English second language subjects by adding data of test scores other than the Peabody Picture Vocabulary Test given in November.

Discussion

The finding that Italian first language children performed at average levels on achievement tests in grade one forces rejection of the idea that immigrant children will always do poorly in school. When factors such as socioeconomic level, nonverbal ability, and comfort within the school system are controlled, English second language children seem to rapidly overcome verbal deficiencies. It should be noted that since the Italian children represented the majority group in each class, the teachers geared the academic programmes to these children. The teachers in this study reported that there was an emphasis on learning English language skills

TABLE 3
STEP-WISE REGRESSION ANALYSIS FOR THE
ENGLISH SECOND LANGUAGE GROUP

Achievement Subtests	Predictor Variables	Multiple R	R ² Change
Word Knowledge	Peabody-Nov.	.568	.322
	Raven	.575	.008
	Peabody-May	.575	.001
	Difference (Nov.-May)*		
Word Analysis	Peabody-Nov.	.403	.163
	Raven	.416	.010
	Peabody-May	.427	.009
	Difference (Nov.-May)*		
Reading	Peabody-Nov.	.401	.161
	Raven	.408	.006
	Peabody-May	.410	.001
	Difference (Nov.-May)*		
Total Reading	Peabody-Nov.	.521	.271
	Difference (Nov.-May)	.522	.001
	Peabody-May*		
	Raven*		
Total Mathematics	Peabody-Nov.	.542	.294
	Raven	.584	.047
	Difference (Nov.-May)	.605	.025
	Peabody-May*		

* *F* level insufficient for further computation.

which seems to have resulted in balancing the English and Italian known by the experimental group. It would be of interest to study the programmes of these teachers since no formal second language curriculum plan existed in these schools.

With regard to the prediction of achievement in English second language children, this study provides clear support for the correlation of English verbal ability and achievement and some support for the relationship of nonverbal abilities and achievement. These findings are consistent with the work of Cooper (1958), Phillipus (1968), and Rattan and MacArthur (1968), which utilized older subjects of varying ethnic backgrounds.

The most dramatic finding seems to be the strong correlation of the Peabody Vocabulary Test administered in November and achievement in the English second language group. It is possible to speculate that English verbal skills acquired prior to kindergarten may be instrumental in academic achievement. A study of the factors which lead to the early acquisition of English by some Italian children seems warranted. Variables such as parental support, older siblings who are learning English, and nursery school exposure to English should be investigated.

The findings of the present study regarding prediction of achievement for monolingual control subjects are somewhat inconsistent with the literature, particularly for verbal skills. The lack of significant correlations between Peabody scores and achievement may be attributed to the small size and somewhat unique characteristics of the control group utilized in this study. These characteristics

include the fact that the monolingual children represented a minority group in the selected schools and had academic programmes which were not directed toward optimal development of their reading and mathematics skills.

This paper is based on a master's thesis completed in the Department of Education, Concordia University. The research was partially supported by a grant from Formation de Chercheurs et d'Action Concertée (Minister of Education, Province of Quebec).

References

- Ashby, B., Morrison, A., & Butcher, H. J. The abilities and attainments of immigrant children. *Research in Education*, 1970, 4, 73-80.
- Bhatnagar, J. *Immigrants at school*. London: Cornmarket Press, 1970.
- Cohen, D. K. Immigrants and the schools. *Review of Educational Research*, 1970, 40, 13-27.
- Cooper, J. G. Predicting school achievement for bilingual pupils. *Journal of Educational Psychology*, 1958, 49, 31-36.
- Dokecki, P. R., Frede, M. C., & Goutney, D. B. Criterion, construct, and predictive validities of the Wechsler Preschool and Primary Scale of Intelligence. *Proceedings of the 77th Annual Convention of the American Psychological Association*, 1969, 4, 505-506.
- Dudek, S. Z., Goldbert, J. S., Lester, E. P., & Harris, B. R. The validity of cognitive, perceptual-motor and personality variables for prediction of achievement in grade I and grade II. *Journal of Clinical Psychology*, 1969, 25, 165-170.
- Dunn, L. M. *Expanded Manual for the Peabody Picture Vocabulary Test*. Minneapolis: American Guidance Services, 1965.
- Durost, N., Bixler, H., Wrightstone, J. W., Prescott, G. A., & Balow, I. H. *Metropolitan Achievement Tests, Primary I, Forms F, G, and H*. New York: Harcourt, Brace, and Jovanovich, 1971.
- Henderson, N. B., Butler, B. V., & Goffeney, B. Effectiveness of the WISC and Bender-Gestalt Test in predicting arithmetic and reading achievement for white and nonwhite children. *Journal of Clinical Psychology*, 1969, 25, 268-271.
- Hirshoren, A. A comparison of the predictive validity of the Revised Stanford-Binet Intelligence Scale and the Illinois Test of Psycholinguistic Abilities. *Exceptional Children*, 1969, 35, 517-521.
- Lesser, K., Schoeninger, D. W., & Bridges, J. S. Prediction of first grade performance. *Perceptual and Motor Skills*, 1970, 31, 751-756.
- Little, A., Mabey, C., & Whitaker, G. The education of immigrant pupils in inner London Primary Schools. *Race*, 1967-68, 9, 439-452.
- McFie, J., & Thompson, J. A. Intellectual abilities of immigrant children. *British Journal of Educational Psychology*, 1970, 40, 348-351.
- Meyers, C. E., Attwell, A. A., & Orpet, R. E. Prediction of fifth grade achievement from kindergarten test and rating data. *Educational and Psychological Measurement*, 1968, 28, 457-463.
- Mussen, P., Dean S., & Rosenbert, M. Some further evidence on the validity of the WISC. *Journal of Consulting Psychology*, 1952, 16, 410-411.
- Nie, N., Bent, D., & Hull, C. H. *Statistical Package for the Social Sciences*. New York: McGraw Hill, 1970.
- Panther, E. E. Prediction of first grade reading achievement. *Elementary School Journal*, 1967, 68, 44-48.
- Phillipus, M. J. Test prediction of school success of bilingual Hispanoamerican children, *ERIC Report # ED 036577*, 1968.
- Plant, W. T., & Southern, M. L. First grade reading achievement predicted from WPPSI and other scores obtained 18 months earlier. *Proceedings of the 76th Annual Convention of the American Psychological Association*, 1968, 3, 593-594.
- Rattan, M. S., & MacArthur, R. S. Longitudinal prediction of school achievement for Metis and Eskimo pupils. *Alberta Journal of Educational Research*, 1968, 14, 37-41.

- Raven, J. C. *Coloured Progressive Matrices, Sets A, Ab, B*. London: Lewis and Co., 1962.
- Raven, J. C. *Guide to using the Coloured Progressive Matrices, Sets A, Ab, B*. London: Lewis and Co., 1965.
- Toronto Board of Education. *Immigrants and their education*. Toronto Board of Education, 1965.
- White, D., & Panunto, B. Verbal and nonverbal abilities in English first and second language children, *Psychological Reports*, 1978, 42, 191-197.

JANA M. MASON

University of Illinois at Urbana-Champaign
and

JANET ROSS KENDALL

Simon Fraser University

Facilitating Reading Comprehension Through Text Structure Manipulation

Two experiments were conducted using 9 to 12-year-old readers to determine how they differ from adults in their ability to identify meaningful units in a written passage, how they are affected by sentences of increasing difficulty, and whether they are aided by either shorter sentences or meaningful segmentation of sentences. The first experiment showed that children mark insufficient or inappropriate intrasentence pausal breaks. The second experiment determined that shortened or segmented sentences may increase reading time and, with low ability readers, will improve reading comprehension. There was also an indication that for all children, reading comprehension can be affected by text characteristics even when sentence length is controlled and word length is manipulated systematically. (Dr. Mason is Assistant Professor, Center for the Study of Reading at the University of Illinois at Urbana-Champaign; Dr. Kendall is Assistant Professor in the Faculty of Education at Simon Fraser University.)

There is evidence that an ability to identify intrasentence units of text is an important aspect of comprehension (Rode, 1974-75; Weinstein & Rabinovitch, 1971) rather than of reading efficiency alone (Coleman & Kim, 1961). Investigators (Okan, Wiener & Cromer, 1971; Steiner, Wiener & Cromer, 1971) have demonstrated that a distinguishing characteristic of good and poor readers is poor readers' inability to identify and organize these units; thus poor readers' comprehension is significantly inferior to that of good readers. This topic should be pursued, both to learn under what circumstances poor readers are impaired and to determine whether manipulations of syntactic complexity can improve reading comprehension.

A technique for identifying meaningful units in written text was first employed by Johnson (1970) who found a high agreement among skilled readers when they were asked to place slash marks at the points where they would pause when reading a passage aloud. Frase and Schwartz (1977) relied on this procedure to show that separation of sentences into indented, meaningful units reduced the length of time

needed to locate information. Hartley and Burnhill (1976), Cromer (1970), Coleman and Kim (1961), and North and Jenkins (1951) found that text materials that were separated into phrases reduced errors in finding information, improved comprehension scores for poor readers, or led to faster reading.

In research with children, Buswell (1920) found that good comprehenders had longer eye-voice spans than did poor comprehenders. This was confirmed by Levin and Kaplan (1970). Eagan (1973) showed that the number of oral reading pauses decreased from grade 2 to 3 and from poor to good readers even after controlling for decoding skill. These studies indicate that children who are better comprehenders have the longer eye-voice spans, implying an ability to identify and group appropriate intrasentence units.

Based on the eye-voice span research, there is reason to expect age and reading skill differences in an ability to identify intrasentence pausal junctures. The effect would be more important, however, if an improvement in reading comprehension could be obtained from text that was segmented by intrasentence junctures, especially if poor readers in the upper elementary grades should obtain higher reading comprehension scores when sentences are segmented into meaningful units than when left unchanged. That is, finding a greater difference between low and high ability readers on the standard text than on the segmented text, with only low ability readers being helped by the manipulation, would indicate that comprehension can be affected by the *obviousness* of intrasentence units.

In the first experiment reported here, adults and 9- to 11-year-old children were asked to read and segment into pausal junctures a passage which was graded by paragraph from a grade 1 to a grade 9 level of difficulty. This study was intended to determine whether children are able to locate intrasentence junctures. In the second study, fourth graders read a story in a standard or revised format. Children marked the time it took them to read each of four sections of the passage; after reading they answered comprehension questions about the passage. Format effects were expected using children's reading time and comprehension error rate as dependent measures; format effects were also predicted as a function of reading ability.

Experiment 1

The purpose was to determine whether children who could recognize most of the words in a text would be able to segment sentences into meaningful units in a manner similar to that of adults.

Method

Twenty-two college students and 60 children aged 9, 10, and 11 read a nine-paragraph passage. The passage had been normed as an oral reading task; the first paragraph was at a grade 1 level of difficulty and each succeeding paragraph increased in difficulty up to grade 9. Adults and children were asked to place slash marks where they would pause if they were to read the passage aloud; children were also asked to circle words they could not pronounce and stop reading if they circled five or more words in a paragraph. The tasks were carried out in classroom settings. Eight children's responses were omitted from the analysis because they circled several words from or did not complete the 5th or 6th grade level paragraph. Of the remaining children, 8 were 9 years old, 21 were 10, and 23 were 11. All of the children appeared to have made realistic assessments of their ability to read the materials.

Only three of the nine paragraphs were analyzed — those approximating the difficulty levels of grades 4, 5, and 6. The easier passages were not analyzed because adults determined that these sentences contained very few breaks (only 3 in the 21 sentences). The more difficult passages were not analyzed because fewer children completed them and those who did tended to restrict themselves to marking commas. In the three analyzed paragraphs there were 22 intrasentence junctures that were agreed upon by half or more of the college students. These junctures were compared with those produced by the children in terms of hits (slash marks that agreed with adults'), misses (slash marks that disagreed with adults' marks), and proportion of responders who parsed only punctuation marks.

Results

Large differences between adults' and children's judgments were found, particularly in the number of slashes. Children marked only six to nine of the junctures that at least half of the adults had located. Smaller differences appeared in the number of misses primarily because some children relied almost entirely on comma junctures for placement of slashes. Forty-two and 43% of the 10- and 11-year-olds but none of the 9-year-olds and adults marked only commas. With these materials, the 9-year-olds made idiosyncratic judgments, ignoring most commas as well as other appropriate junctures; 10- and 11-year-olds were quite cautious in their judgments; and adults, while agreeing on the essential pausal breaks, tended in addition to mark lesser pause breaks (see Table 1 for summary).

TABLE 1
TABULATION OF HITS, MISSES, AND AN ACCURACY RATIO
OF INTRASENTENCE PAUSAL BREAKS

Age	Average hits	S.D.	Average misses	S.D.	Hits/ Total response
9	6.5	3.1	8.1	7.9	.44
10	9.0	4.8	3.4	6.5	.73
11	9.2	4.2	3.0	4.9	.75
College	15.3	5.1	6.6	6.3	.70

Discussion

Children cannot readily separate sentences into meaningful units although the ability to do so shows some improvement with age. Nine-year-olds make little use of any consistent rule in marking boundaries; however, it should be noted that two of the passages may have been too difficult for them to read. The greater restraint and accuracy of 10- and 11-year-olds over the younger children and even adults indicates that although the task is understood, there is an insufficient ability to locate pausal boundaries.

It should be noted that because only eight 9-year-olds remained in the sample, differences between them and older children must be interpreted cautiously. The major point, that children make judgments quite different from those of adults, is unaffected by the small number of 9-year-olds.

Experiment 2

An inability of upper elementary school children to distinguish in sufficient detail intrasentence units might be related to reading comprehension. In the second experiment, this matter is tested by hypothesizing that low ability readers cannot locate intrasentence junctures in text materials but if junctures are made more obvious, comprehension scores ought to be improved and reading rate changed. This is tested by measuring the reading comprehension and reading rate effects of texts materials which are changed specifically to make intrasentence units more apparent.

Method

Subjects. Ninety-eight fourth grade students from a metropolitan Canadian city who had not been used in the first experiment were tested. Schools and classrooms were selected on the basis of Metropolitan Achievement Test (MAT) score averages which were at or near the 50th percentile.

Materials. The three paragraphs that were analyzed in Experiment 1 served as a content base for the construction of systematically varied materials. Four levels of difficulty were created by dividing the paragraphs into 4 sections, each of which contained five sentences; each sentence in the first section contained 18 syllables, the next had 22 per sentence, then 26, and the fourth section had 30. Five multiple-choice comprehension questions were constructed for each of these four levels. Two variations of these materials were put together so that in one, pausal units appeared on separate lines of the page while in the other, four short sentences reflecting these units replaced each original sentence. The complete structure is summarized in Table 2. An example of the materials is shown below.

Standard

Dick will be in Grade Five and though he enjoys math he likes art class best.

Parsed

Dick will be in Grade Five
and though he enjoys math
he likes art class best.

Short sentence

Dick will be in Grade Five. He enjoys math. He enjoys art. He likes art class best.

Questions from comprehension test

What is Dick's best subject?
(a) math (b) reading (c) art (d) science

Procedure. In group testing, children read the passage in one of three format conditions: (1) standard text, (2) each sentence broken into four short sentences, and (3) each sentence divided by meaningful units onto separate indented lines. When they finished each of the four difficulty level sections, they wrote down the time that was displayed on the blackboard (the time was changed every 15 seconds). When they finished the passage, they responded to the vocabulary and comprehension questions which followed. They were instructed to answer the

TABLE 2
DESCRIPTION OF FORMAT AND INTRAPASSAGE DIFFERENCES

Format condition	Passage section			
	First	Second	Third	Fourth
Short sentence				
Number of lines	9	11	14	15
Number of sentences	20	20	20	20
Number of words per sentence	5.35	7.25	9.30	10.15
Number of syllables per sentence	5.85	5.85	6.40	6.55
Parsed				
Number of lines	11	12	11	12
Number of sentences	5	5	5	5
Number of words per sentence	17	17	17	17
Number of syllables per sentence	18	22	26	30
Standard				
Number of lines	8	8	9	9
Number of sentences	5	5	5	5
Number of words per sentence	17	17	17	17
Number of syllables per sentence	18	22	26	30

questions without rereading the story and were given as much time as they needed to read and to answer the questions.

Results

Before running an analysis of variance, scores from the *MAT* were used to construct 3 approximately equal-sized ability groups (the low group scored between 5 and 20 and the high group was between 30 and 42). Then, *passage format* and *reading ability* were between subjects variables while *passage difficulty* was a within subjects variable. This design was analyzed three times using comprehension error, reading time, and the ratio of reading time over the number of syllables per standard sentence as dependent variables. Mean scores from these analyses appear in Table 3.

Comprehension error. With an average error rate of 28%, significant effects were obtained of reading ability, $F(2,87) = 15.0, p < .001$, passage difficulty, $F(3,261) = 23.0, p < .001$, and a borderline interaction between reading ability and passage format, $F(4,87) = 2.4, p < .06$. However, an a priori test was warranted in analyzing the interaction because low ability readers had been predicted to be helped by the experimental formats. *T*-test comparisons of the format conditions indicated significant differences for low ability readers between the standard and short sentence formats, $t(87) = 4.29, p < .01$, and between the standard and parsed sentence formats, $t(87) = 2.0, p < .05$. No significant *t*-test differences were found for middle or high ability students. The interaction is shown in Figure 1, on the left.

The main effect of passage difficulty indicated, contrary to expectation, that paragraphs containing the shortest words (first section) and longest words (last section) generated the lowest error rates. This will be discussed later.

Reading time. The reading time analysis showed significant effects of reading

TABLE 3
ERROR RATES AND READING TIMES UNDER PASSAGE FORMAT AND
PASSAGE DIFFICULTY CONDITIONS, AND AS A FUNCTION OF
READING ABILITY

	Mean		Ratio of Reading time to Number of syllables
	Error rate	Reading time in seconds	
Format			
Standard format	1.51	39.7	.34
Parsed format	1.46	45.3	.38
Short sentence format	1.29	55.5	.34
Passage difficulty (section)			
First	1.00	34.2	.35
Second	1.72	47.1	.39
Third	1.94	50.3	.34
Fourth	1.02	55.6	.33
Reading ability			
High	1.04	38.3	.29
Middle	1.27	42.6	.32
Low	1.95	59.5	.44

ability, $F(2,87) = 7.1, p < .001$, passage difficulty, $F(2,87) = 33.9, p < .001$, passage format $F(2,87) = 3.6, p < .03$, and a passage format by passage difficulty interaction, $F(6,261) = 3.0, p < .01$. The easiest passage section took an average of 32 seconds to read, the hardest took 56 seconds. Reading ability (MAT scores) was correlated significantly with reading time under the two experimental conditions at every level of difficulty, but not with the standard format (see Table 4).

TABLE 4
CORRELATIONS BETWEEN MAT COMPREHENSION AND READING TIME
AS A FUNCTION OF TEXT FORMAT AND TEXT DIFFICULTY

Format	Text Difficulty				Whole Text
	1 (Easiest)	2	3	4 (Hardest)	
Standard	+.01	+.07	+.04	-.20	-.05
Parsed	-.61*	-.59*	-.53*	-.60*	-.64*
Short sentence	-.58*	-.54*	-.55*	-.59*	-.60*

* $p < .01$

The reading ability effect indicated that high ability readers were the fastest readers while low ability readers were the slowest. These differences were maintained over all levels of text difficulty but not uniformly over format conditions. Although the latter interaction was not significant, since a difference had been predicted and correlations indicated a discrepancy, Tukey's test was made to evaluate the effects of format on reading time of high and low ability

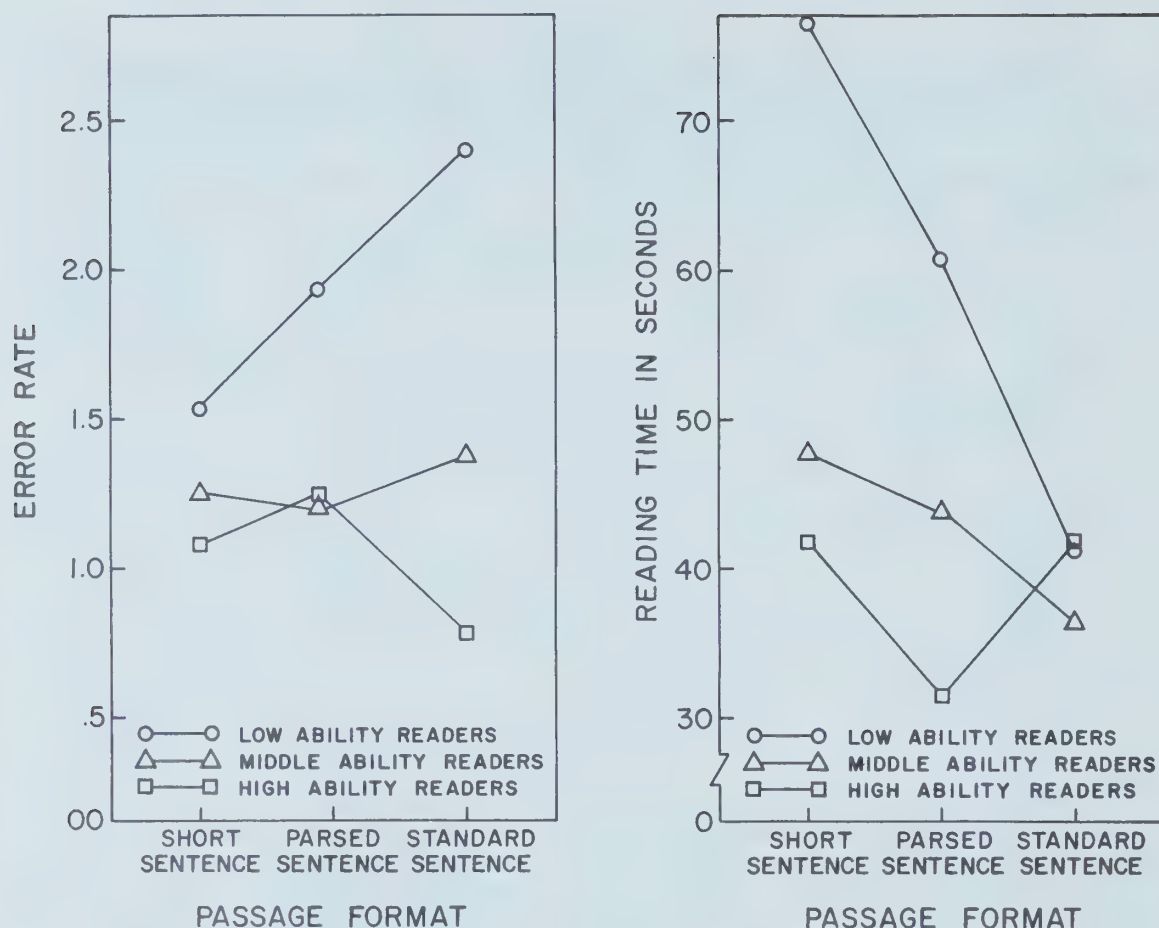


Figure 1. Interactions between reading ability and passage format for error rate and for reading time.

readers. There was a significant difference under the short sentence condition, $q(3,32) = 4.2$, $p < .05$, a nearly significant difference in the parsed condition, $q(3,32) = 3.1$, $p < .10$, and no difference in the standard condition. These effects are graphed on the right side of Figure 1.

The format main effect determined that the experimental formats increased reading time with the short sentence taking the longest time. The interaction between passage format and passage difficulty showed that as the number of syllables per sentence was increased, more time was spent reading. This was particularly true for the two experimental formats.

In a third analysis using a ratio of time over syllable number, there were neither significant format differences nor a significant interaction with passage difficulty, permitting the interpretation that format differences where time alone was the dependent measure were caused by the short sentence format which contained fewer words per sentence. The passage difficulty effect was still significant in the ratio analysis, $F(3,261) = 5.6$, $p < .001$, as was reading ability, $F(2,87) = 6.7$, $p < .002$.

Frase and Schwartz (1977) found that appropriately parsed and indented sentences reduced reading time. In this study, parsing did not reduce reading time. This is probably due to the task differences. Frase and Schwartz had people read until information was found rather than read to understand and remember information. Also, the amount of text read was longer in the former study. If

children had been given a longer passage to read, their reading speed might have been lower under parsed format conditions than under the standard format.

Discussion

The error analysis determined that formats which simplify the syntax can improve reading comprehension for low ability readers but not for better readers. Low ability readers' error rates are reduced under both short sentence and parsed sentence formats. The source of the improvement appears to be explained partially by the amount of time spent on the passage. Although subject to individual variation, most low ability readers spend more time than do better readers on the nonstandard format text but the same time as others on regular text. This suggests that low ability readers ignore the greater complexity of or do not parse properly the more complicated intrasentence units. This supposition is supported by Weinstein and Rabinovitch (1971) who found that, in contrast to high ability readers, low ability readers do not improve in the number of nonsense words learned when syntactic cues are provided. The interpretation is further supported by the effects of the experimental formats and the correlations with reading ability (Figure 1 and Table 4). When intrasentence units are set out on separate lines or placed in separate sentences, low ability readers spend appropriately more time on the task and obtain comprehension scores that are closer to those obtained by better readers. High and middle ability readers do not spend more time on the experimental formats and are not helped by manipulated text. This suggests that better readers understand and locate while reading appropriate intrasentence units of text materials used here. It is possible that with text structures that are more complex than these, better readers could be helped by a parsed format; however, this remains to be shown.

Comprehension errors are not distributed in a linear fashion even when sentence length is held constant and word difficulty is increased systematically. The unexpected ease of the fourth section is interesting, particularly in the light of assertions that word length and sentence length are sufficient variables to measure text difficulty (according to commonly used readability formulas). One explanation of the low error rate on the passage section which contains the longer words (30 syllables per sentence) is that variations in text cohesiveness affect text difficulty (Stein & Glenn, 1977). The second and third sections consist of poorly related pieces of information (list-like structures) while the first and last sections have story-like structures. Since the comprehension question on all sections are of a factual nature and thus are of approximately equal difficulty, it can be concluded that scores on multiple choice tests of comprehension are affected by text cohesiveness as well as by word length.

Summary

The first experiment demonstrates that children are not able typically to distinguish meaningful intrasentence units, particularly as sentences increase in length and contain longer words. Intrasentence junctures are inadequately identified, even by sixth graders, when compared with adults' judgments.

The second experiment indicates that only low ability readers' comprehension scores are improved when intrasentence units are separated. It occurs in conjunction with a reduced reading rate, perhaps because, under conditions when the syntax is more apparent, the low ability readers now process the more complex

information. It is conceivable, then, that separating sentences into parsed structures may help low ability readers cope with the syntactically more complex text materials that prevail from fourth grade onward.

Several questions remain unanswered. Will the effect be stronger: When longer sentences are given? When time constraints are imposed? If longer passages are used? If other measures of reading comprehension are given? Obviously there are a number of studies that should be undertaken before effects on comprehension of segmented sentences are fully understood.

Additionally, the fact that children do not agree with adults' conceptions of pausal junctures (Experiment 1) means that they might be trained to locate more of the necessary intrasentence boundaries. It is possible that this kind of task could improve comprehension of complex sentences. This effect needs to be demonstrated by a training study.

Nonconfirmation of a linear fit between comprehension error rate and word length is also important. Text complexity is presently based on only two factors, word length and sentence length. It is apparent that other variables such as story cohesiveness enter into text difficulty. As story grammars are more fully developed, the role in comprehension of text-level structures can be explored. This study suggests that text cohesiveness is an important third variable in assessing text readability.

References

- Buswell, G. An experimental study of the eye-voice span in reading. *Supplementary Educational Monographs* #17. Chicago: University of Chicago Press, 1920.
- Coleman, E., & Kim, I. Comparison of several styles of typography in English. *Journal of Applied Psychology*, 1961, 45, 262-267.
- Cromer, W. The difference model: A new explanation for some reading difficulties. *Journal of Educational Psychology*, 1970, 61, 471-483.
- Eagan, R. *An investigation into the relationship of the parsing phenomena in oral reading and reading comprehension*. Unpublished doctoral dissertation, University of Alberta, 1973.
- Frase, L., & Schwartz, B. *Typographical cues that facilitate comprehension*. Murray Hill, N.J.: Bell Laboratories, 1977.
- Hartley, J., & Brunhill, P. Explorations in space: A critique of the typography of BPS publications. *Bulletin of the British Psychological Society*, 1976, 29, 97-107.
- Johnson, R. Recall of prose as a function of the structural importance of the linguistic units. *Journal of Verbal Learning and Verbal Behavior*, 1970, 9, 12-20.
- Levin, H., & Kaplan, E. Grammatical structure and reading. In H. Levin and J. Williams (Eds.), *Basic studies on reading*. New York: Basic Books, 1970.
- North, A., & Jenkins, L. Reading speed and comprehension as a function of typography. *Journal of Applied Psychology*, 1951, 35, 225-228.
- Oakan, R., Wiener, M., & Cromer, W. Identification, organization, and reading comprehension for good and poor readers. *Journal of Educational Psychology*, 1971, 62, 71-78.
- Rode, S. Development of phrase and clause boundary reading in children. *Reading Research Quarterly*, 1974-75, 10, 124-141.
- Stein, N., & Glenn, C. An analysis of story comprehension in elementary school children. In R. Freedle (Ed.), *Multidisciplinary approaches to discourse comprehension*. Hillsdale, N.J.: Ablex, Inc., 1977.
- Steiner, R., Wiener, M., & Cromer, W. Comprehension training and identification for poor and good readers. *Journal of Educational Psychology*, 1971, 62, 506-513.
- Weinstein, R., & Rabinovitch, M. Sentence structure and retention in good and poor readers. *Journal of Educational Psychology*, 1971, 62, 25-30.

MARILYN ASSHETON-SMITH

The University of Alberta

and

KELLEEN TOOHEY

Ontario Institute for Studies in Education

School-Centered Community Conflict: The Holdeman Mennonite Case in Alberta*

The Holdeman Mennonites, or Church of God in Christ Mennonite, in 1976 began to establish private schools in eight rural communities across Alberta. This decision generated varying degrees of resistance from other community members, and varying amounts of conflict. This paper provides background information on the dispute, documents the conflict in two affected farming communities, and seeks to explain the difference in conflict through analysis of ideological and structural differences in the two communities. (Ms. Assheton-Smith is an Assistant Professor and Coordinator of the Intercultural Program in the Department of Educational Foundations at the University of Alberta; Ms. Toohey is presently studying towards her Ph.D. in Education at the Ontario Institute for Studies in Education.)

The school is frequently the focus of community-based conflict. This conflict is often related to social divisions within the community, but the existence of social cleavage does not always generate conflict. Theoretical explanations for the presence of conflict in some cases and its absence in others continue to be only partially satisfactory. (See, for example, Gamson, 1966; Lee & Lapointe, 1977).

In Alberta, in 1977-78, a case of school-centered conflict occurred which represented a "natural experiment" for the study of community conflict, since in two closely comparable communities the same decision generated high levels of conflict in one and low levels of conflict in the other. The case involved the decision of Holdeman Mennonite parents across Alberta to withdraw their children from the local public schools and to establish independent Holdeman-controlled schools.

*The authors would like to acknowledge gratefully the assistance of the Church of God Mennonite parents, school board members in both Camrose County and Spirit River School Division, as well as past and present members of the public school board and superintendents in both jurisdictions. The research was partially supported by an Operating Research Grant from the General Research Fund of the University of Alberta.

This study will first document the situation in the two communities, demonstrating thereby that the level of school-centered conflict was indeed high in one community and low in the other. It will secondly seek to explain why there might have been more conflict created by the Holdeman decision in one community than in the other. For explanatory purposes, two theoretical orientations will be drawn upon.

The first theoretical orientation is derived from Carlton (1974), who states that different images of the school held by different sectors of the public may explain school-centered conflict in a given situation. Carlton argues that "an understanding of the way in which people think about the school is relevant to their behaviour in relation to the school" (1974, p. 73) and that at times some images of schooling have achieved very wide acceptance, becoming important influences in the development of school systems.

The description of a relationship between thoughts, or images of schools, and actions implies a kind of idea system which Guy Rocher would term an ideology. These ideas have, in Rocher's terms, a conative function: "pushing . . . a collectivity to action, or at least directing the collectivity by providing goals and means" (1972, p. 103). Translating Carlton's statements into Rocher's terms,¹ school-centered conflict at the ideological level would be based upon different ideologies of schooling held by different groups.²

The second theoretical orientation is generated by adding to Carlton's notions a perspective provided by Gerth and Mills (1953)³ and the general sociological orientation given by Coulson and Riddel (1970). Carlton views his school images, or school-related ideologies, as dependent upon the social institution to which the school is believed to be responsible. These include, among others, the family, the state, religious groups, and professional teachers. Gerth and Mills (1953, p. 251), dealing with a very similar set of social institutions, would convert Carlton's images to theoretical types with the argument that schools have little autonomy from other social institutions. That is, it is not just a matter of beliefs, because schools are simply extensions of one or another of the social institutions named by Carlton. If this is the case, conflict around schools should not be conceptualized merely as a conflict of "images," but rather as conflict relative to which of the other social institutions will, in fact, control schooling and thereby determine the nature of schools.

Coulson and Riddel potentially extend this type of conflict analysis by treating social analysis as the process of analyzing groups in interaction within a social structure which benefits some groups more than others. Conflict arises when one group seeks to improve the benefits to itself and another group recognizes this as a threat. The details of the conflict depend upon the consciousness or understanding of the situation which each of the participant groups have. But the analyst must seek to understand both the manifest explanations of the conflict, which are ideologically related explanations given by the participants themselves, and the latent explanations which are the causes embedded in the social structure. The actors themselves may not be aware of these structural causes.

Methodology and Data Collection

The methodology of this research consists of a comparative case study approach, with two cases selected on the basis of variation on one dependent variable: the level of conflict created by the Holdemans' decision to establish

independent schools. Both communities are described structurally and documentation of the school-related beliefs (or ideologies) of Mennonites, professional teachers, and school board members of Camrose County and Spirit River Division is presented. The authors then consider whether ideological differences or structural differences account more adequately for the observed differences in the dependent variable.

The data collection procedures were two-fold. First, the research team conducted unstructured interviews with persons who were members of each of the categories outlined above (Mennonites, professional teachers, and school board members). These were carried out in the homes or offices of the respondents. The interviewers sought both to obtain information about the events surrounding the establishment of the independent schools and to obtain information on the beliefs about schooling held by the respondents. The second data collection procedure consisted of the examination of available data from county and school division offices, and from the Department of Education in Edmonton. These included documents such as annual reports, enrollment records, financial records, and county maps.

Neither the cases selected nor the respondents interviewed should be considered representative as that term is used in probability sampling. Rather, all were selected with the above kinds of theoretical problems in mind. At present, inferences must be drawn with caution because of the exploratory nature of the research, and generalization to other similar cases should also be made with care.

Background Information: The Holdemans and Alberta Schools

The Holdeman Mennonites are one of a number of subdivisions of the followers of Menno Simons. Their historical origins were in sixteenth century religious conflict in Central Europe and involved both a rejection of the Roman Catholic authority and rejection of involvement with the civil authority. They could be seen then as early proponents of the separation of church and state, but also as early practitioners of adult-shared community decision-making, based upon biblical teachings rather than on response to an hierarchical authority structure.

A number of Mennonites migrated to Canada beginning in 1874 because of the availability of land in Manitoba and the assurances of the Canadian government that they would have freedom from military service and the right to their own schools. One Joseph Holdeman, himself originating from a different migration of Mennonites to the United States, began preaching a conservative revivalism in 1858; the Alberta Holdeman group owe their intellectual origin and contemporary conference membership to this branch of the Mennonite faith.

The Holdemans began their settlement in Alberta at Linden, near Three Hills, in 1902. In the 1930s, some families headed north to settle at Crooked Creek in the Peace River country and in the early 1960s, families began to move into the Spirit River area and into Camrose County. Congregations currently exist in these as well as in the Stettler and Pincher Creek districts. The total number of persons involved is approximately 800, representing a small fraction of the total Mennonite population in the province of Alberta.

The Holdemans began to establish their private schools in Alberta in 1976. At the time of the research, the two Holdeman schools in Spirit River School Division were in their second year of operation and those in Camrose County in their first

year. By 1978, there had been a total of eight such schools established in the province with, at most, 300 children in attendance.

The 1978 court case in Alberta involved essentially an attempt to return Holdeman children to the public school at Linden (near Three Hills) through legal intervention. Charges were laid against forty-five Holdeman parents (with one parent being the test defendant) because their children were truant from a school approved by the Minister of Education. The Holdeman Schools were not approved because they did not have certified teachers and had not been inspected by school authorities. In a decision handed down on February 6, 1978, Judge Oliver ruled that the test defendant was not guilty, by virtue of the argument (among others) that specifically, in the case of the Holdemans, the School Act conflicted with Section 2 of the Alberta Bill of Rights guaranteeing religious freedom. During the first week of March, 1978, Julian Koziak, the Minister of Education, announced that a fourth type of approved school would be added to the Alberta School Act. This type of school could operate without certified teachers, but would receive no funding from the government. This decision removed, at least for the time being, the legal threat to the Holdeman schools.

Description of the Conflict

As noted earlier, two of the Alberta school jurisdictions which experienced Mennonite withdrawal of students seemed to respond quite differently. In Spirit River School Division, although board members were concerned about the Mennonite decision and participated in meetings with other boards and the Department of Education to discuss the situation, they stated publicly that the Holdeman Mennonite schools seemed adequate. In Camrose County, fear of the school system being destroyed was expressed and parents petitioned against the Mennonite school being allowed to continue.

At the local community level, the difference in intensity of response or reaction was most marked. At Hillstream in Camrose County, the Home and School Association circulated a petition against Mennonite schools in February of 1977 and, in a matter of days, had gathered sixty signatures. They presented their petition to the County School Board, asking members to read and to sign it. The School Board refused to sign the petition as a Board but agreed that board members could sign it as individuals (*Camrose Canadian*, February 2, 1977, p. 1).

The Home and School meeting in Hillstream in April of 1978 had as its topic, the Mennonite school situation, and as its guests, representatives of the provincial Department of Education and the Alberta Teachers' Association. Fifty persons attended that meeting, a fact which suggests that neither the court decision nor the minister's legislation satisfied community objections to Holdeman schools. By contrast, there was no organized reaction in the Spirit River district.

There was also considerable difference at the local level in the newspaper coverage of the topic. Residents of the Spirit River School Division are served by a number of regional newspapers. As far as could be ascertained, there was no newspaper coverage of the withdrawal of Holdeman Mennonite children from the area's public schools. A regional paper, the *Grande Prairie Herald-Tribune*, began in September of 1977 to carry Canadian Press stories of the court case at Linden but contained no investigation of, or even reference to, the local private Mennonite schools. The smaller papers as well printed no stories of the local issue, nor did

they carry the Canadian Press stories about the court case at Linden. In fact, very few stories relating to schools were carried in these papers.

On the other hand, the regional newspaper in Camrose, the *Camrose Canadian*, routinely reports on County council meetings, and thus on school business. As early as February, 1977, the *Camrose Canadian* was reporting on school board discussion of the Mennonite proposal to withdraw their children. It also printed letters to the editor on the subject and generally covered news of the conflict.

Description of the Two School Jurisdictions

Case A: Camrose County and "Hillstream" School

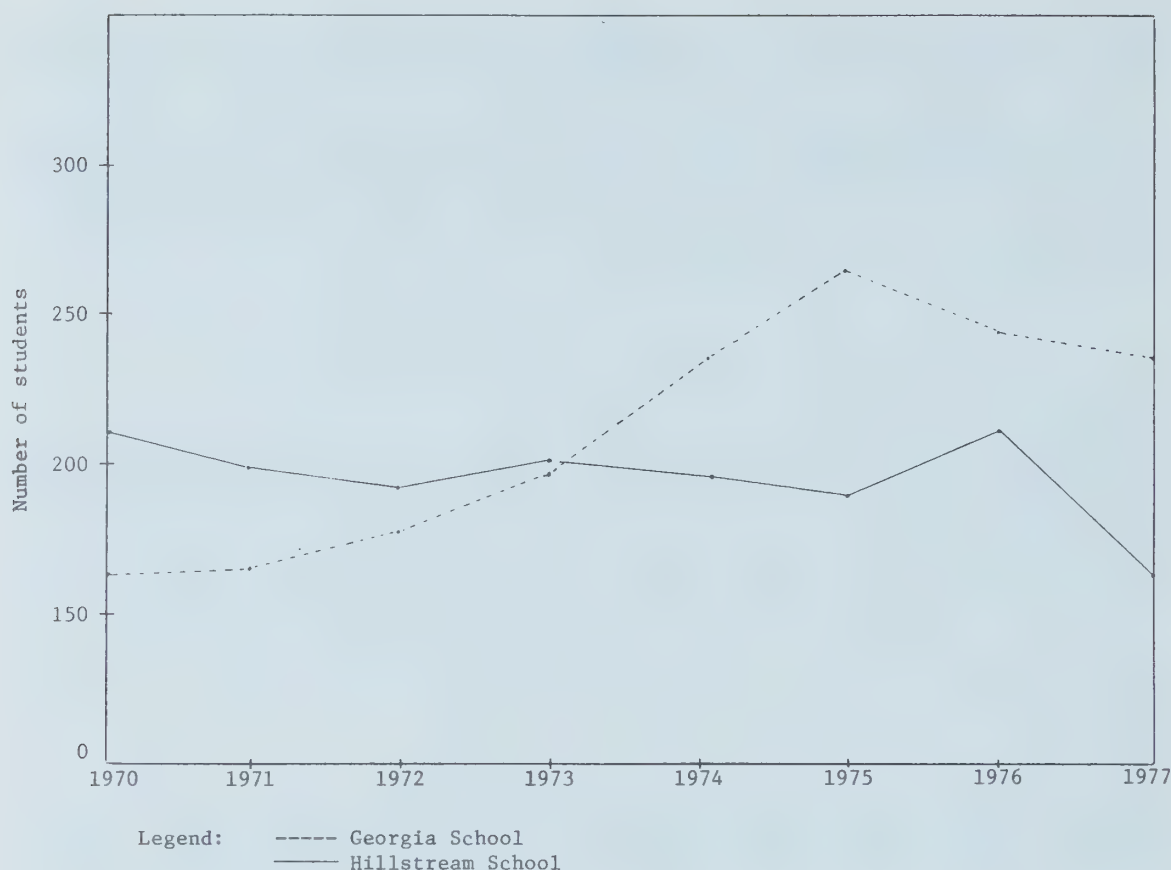
Camrose County is located in central Alberta, some 100 km southeast of Edmonton. County offices are located in Camrose, a city with a population of 11,000, which has its own municipal government and school boards. In addition to the 3,500 square km of farm land under county jurisdiction, there are eight incorporated villages and towns, the schools of which are part of the county system.

School centralization resulted in all country schools being centralized into villages and towns. There is no separate school within the county and, until the Holdeman Mennonite decision, there were no private schools. However, county students may enroll in the Camrose Separate (Roman Catholic) school, or in Camrose City high schools. In 1977, 2,136 students were enrolled in the county system, and 264 county residents were enrolled in other school systems. To these 264 must be added the 55 Mennonite children; that is, 17% of all resident children enrolled outside the county schools are Mennonite. All of these 319 children represent a per capita funding loss to the county in relation to provincial funding of schools.

Hillstream school, from which the Holdemans withdrew their children, is located on the edge of Hillstream village, whose population is about 140; no Holdeman Mennonite families reside in the village. The farm population began to settle in the area on homesteads in 1899; Mennonite families started to purchase farms in 1959. Mennonites now make up about 25% of the ninety farm families in the district, but their children represent 50% of the farm school children because of their larger family size. School enrollment was just over 200 in 1962, declined to 160 in 1968, but jumped to 210 in 1972 with the addition of grades nine to twelve from a nearby village. Enrollment remained fairly stable (near 200) from that time until the withdrawal of the Holdeman children in the fall of 1977 when enrollment dropped to 155, with grades one to nine dropping from 143 to 99 students (see Figure 1).

The number of teachers on staff reached its maximum of 14 in 1976; it was down to 12 in 1978. With five of the teachers working at the high school level, there is some doubling of grades at the elementary level. Of the teachers on staff, half grew up in the county and generally attended schools much like Hillstream, and four were raised in the immediate area. The majority of teachers live in the village or on nearby farms, and one is presently mayor of the village.

The Holdemans established their schools in the fall of 1977. There are two schools, one at the southeast edge of the school enrollment area and one in the northwest corner. In each case, the teacher is a man schooled to about the grade ten level in public school. Also teaching or assisting at each school are women who have completed grade ten or eleven. The schools teach from grades one to nine.



Note. Holdeman children were withdrawn from Georgia School in 1976, from Hillstream School in 1977.

Figure 1. School enrollment, 1970-1977: Hillstream and Georgia schools. (Source: Annual Reports, Camrose County No. 22, 1970-1977; Spirit River School District No. 47, 1970-1977.)

Case B: Spirit River School Division and "Georgia" School

The municipal level of government in this case consists of Municipal Districts and Improvement Districts. The School Division, a separately elected and funded structure, incorporates all of the territory in the Municipal District of Spirit River plus much of the land area of two Improvement Districts. It is located in the Peace River area, some 600 km northwest of Edmonton. The School Division offices are located in the town of Spirit River which has a population of just over 1,000. This town's public schools are within the School Division. There are six incorporated towns and villages in the Spirit River School Division, but ten schools, because some of the schools were centralized in villages and towns whereas others were centralized in farming districts without associated villages.

The total enrollment in the school system in September 1977 was 1,686 students, but an additional 170 resident students attended other schools (Roman Catholic Separate School; Seventh Day Adventist and Mennonite Private Schools, and a nondivisional high school). Of these 170 students, just 28, or 17%, are Holdeman Mennonites — exactly the same percentage as in Camrose County.

The Georgia farming community, unlike Hillstream, has no village within its boundaries, although a store is situated across from the school and a recreation complex was opened recently. These three buildings — the school, store, and

recreation complex — provide a minimal service centre for the surrounding area and as such are roughly equivalent to the village in the previous case. However, the majority of the population of this centre are teachers living in eight mobile homes on the school grounds. The farm community was opened to homesteading in the 1950s, at which time Holdeman Mennonite farmers and others began to move in, clear the land, and settle. There are fewer than twenty Holdeman families in the area, representing less than 10% of the total farm population.

Georgia school was first established as a three-room school in 1965; rooms have been added on four occasions since that date. In 1976, a high school addition was constructed, enabling the first children to start grade one in Georgia school to complete their grade twelve there. School enrollment has shown a general increase, from 60 in 1965 to 230 in 1978. The Holdeman Mennonites withdrew their children from the school in 1976, resulting in an enrollment loss of only 13 children (see Figure 1).

The highest complement of teachers in this school was 15 in 1976, dropping to 13 in 1977 and remaining the same in 1978. Of those teachers, only one is from the Peace River country, the rest being from outside the province and in many cases from outside the country. In the past year or so, there has been some sign of this teacher population stabilizing, with about half purchasing land in the area or building homes and apparently deciding to settle. As such, their recent and diverse origins are not unlike those of their neighbours who also have moved into the area within the last ten or fifteen years.

The Holdeman Mennonite school has one room and offers grades one to nine to thirteen students.

Data on Group Beliefs

The following material represents primarily the data collected through interviews dealing specifically with each group's "school-ideologies" or beliefs about schools and their understanding of the situation about Holdeman withdrawal of students from public schools. In some cases, this has included beliefs about the society and about the actions or beliefs of other groups involved. The data were collected between March 10 and April 10, 1978, prior to a media focus in Alberta on an incident in the Mennonite school at Ridge Valley. Such a media event could potentially have modified some of the responses reported here.

Holdeman Mennonite Beliefs

The Holdeman Mennonites' statements about their beliefs and understanding of the situation were consistent in both communities investigated. Although they did not necessarily object in major ways to the nearby public schools which their children had attended, they did object to the general institution of public schooling. Society is perceived as losing its moral focus and this is mirrored in the public schools. At one time rural schools were somewhat protected from events in urban areas, but this is no longer felt to be the case and the rural schools now have many of the problems of urban schools. The "worldly" task of the Mennonite is to follow scripture; this task is felt to be best carried out on the economic base of farming. Parents and the community cannot ignore their responsibility to guide their children into adult roles.

In contrast to the public schools, Holdeman schools see as their goal teaching children the 5 R's — reading, writing, arithmetic, responsibility, and respect.

School is a place of work; recess is a time for play. Adults who have a proper relationship with God and who are committed to teaching can teach the desired skills and virtues. In fact, the most important criteria for choosing teachers is spiritual righteousness and dedication to teaching. It is felt that children can complete the standard curriculum in eight months, but as there is no desire to create undue tension with neighbours, they are kept in school for almost the standard school year. The Holdeman curriculum follows Alberta requirements in general, although texts with appropriate biblically based discussion are used. The school board members of the eight established Holdeman Mennonite schools in Alberta assist each other in the selection of curriculum material. In Spirit River School Division, it was stated, the public school donated old texts for use in the new private schools.

The need to maintain autonomy from the state was clearly indicated. As one respondent noted, "In Manitoba first they provided the firewood, then they paid for the teachers, then they built the school . . . then they started to tell us how to run it." The Holdemans expressed little interest in the working of government; they pay their taxes but do not attempt to determine where the money goes or how it is used. The above Holdeman belief system was consistent across both school jurisdictions and thus had the nature of a "shared ideology," a well-articulated, systematized belief structure which forms the basis for action or legitimizes action once it is taken.

Professional Teacher Beliefs

Teachers in both jurisdictions were fairly consistent in their beliefs. As a result of the situation, they tended to initiate discussion about school beliefs by emphasizing the necessity of teacher certification for the teaching task. Although it was stated that some persons might be able to teach without university training, the general notion was that certification provided a control of standards and expertise that would ensure the best teaching.

There was also a belief expressed that teacher certification and the structure of supervision associated with public schools would result in the actual following of the Alberta curriculum in the classroom. It would seem that belief related to both the teacher-learned skills associated with curriculum study and the notion that teachers accept the legitimacy of the Alberta school curriculum so they would follow it even if they were not under supervision. Teachers expressed the belief that this curriculum ensures a certain minimum level of education in schools. The extent to which they expected Mennonites to deviate from that course of study was a real concern. They expressed particular concern that science would not be taught in the private schools.

Teacher respondents did not always find the goals of schooling easy to articulate. Although there were references to the need to learn the 3 R's, the desired consequences of schooling expressed were a great deal more complex: developing social skills, acquiring a level of competence which enables students to advance to other educational institutions, being able to make a living and contribute both to society and to the employment sphere.

The only clearly expressed state-related role described by teachers was enforcement of the law by the courts: if the law demands compulsory school attendance, certified teachers in the school, or approval of private schools, the courts should enforce those laws. The general impression of teachers was that the

courts tended not to enforce these laws, even in cases other than the Holdeman dispute. Teachers referred to the problems of spending four years obtaining a university degree, becoming a certified teacher, and then having employment prospects threatened by the development of schools which would use uncertified teachers. The belief was expressed that the school boards would now find ways to use uncertified teachers as less costly personnel, and that eventually independent schools such as those of the Mennonites would seek and obtain government funding.

The above described teacher belief system was consistent across both jurisdictions and like that of the Holdemans, it could be seen as part of a “shared ideology.”

School Board Member Beliefs

All interviewed school board members from both jurisdictions expressed a general feeling of discontent with the public schools as they are, echoing some of the reasons the Mennonites gave for disapproval of public schools in general; in fact, almost all initiated discussion by indicating agreement in some sense with the Mennonites’ analysis and subsequent decision to withdraw their children.

School board members saw teacher certification as an issue of importance primarily to teachers, but were very concerned with the cost of schooling. In some cases, this meant objecting to “costly frills” such as outdoor education; in other cases, it meant objecting to the cost of certified teachers and the teachers’ union.

The source of the problem in this economic squeeze was felt by board members of both school jurisdictions to be related to inadequate funding. In Camrose County, school board members were more likely to be struggling with the notion of reducing costs whereas in Spirit River, they were attempting to persuade the government to recognize the inadequacy of current funding policy for their particular situation.

In 1975, there was a major funding crisis in both school jurisdictions. Both made an attempt to pass bylaws permitting a funding increase, bylaws defeated by the electorate in both cases. In Camrose, this was used as justification to close two schools, but no schools were closed in the Spirit River School Division.

Analysis and Conclusions

Following the interviews and extensive search of local press coverage in each community, it was evident that there actually was a higher level of conflict and resistance to the Mennonite decision in the one community than in the other. However, it was apparent that there was more conflict at the local community level than at the larger school jurisdiction level. In the one case, an active Home and School Association provided a structure for expressing resistance while in the other case, there was no such structure nor evidence of resistance. At the county or school division level, the major and, at times, desperate problems of school financing seemed to override any emphasis on the Mennonite withdrawal itself. This was the case particularly in Spirit River.

Why would there be such a difference in the degree of conflict between the two cases? Was it strictly ideological, a conflict between opposing belief systems, or were there significant structural differences between the two cases?

At the ideological level, the Holdeman Mennonite beliefs were totally consistent

in both communities, and consistent with news reports of their beliefs and motivations elsewhere in the province. This consistency tended to be interpreted by others as evidence of obedience to an authority structure but it is also consistent with a fully articulated shared belief system which is held by Holdemans in Alberta and elsewhere. Provincial, national, and international conferences (such as the one in Kansas in 1973 which apparently sparked the schools' decision) would provide settings for exchange of information, world-views or ideologies, and exchange of proposals for action appropriate to the beliefs.

Similarly, the teachers were, on the whole, consistent in their presentation of a professional ideology. Their belief system includes not only the notion of teaching expertise and standards linked to certification, but also some ideas about curriculum which, they appear to believe, should be scientific. They expressed opposition to the "religious" curriculum they believed the Mennonite schools would use, but generally believed that the public schools could accommodate the Mennonites and that it should not have been necessary for the Holdemans to withdraw their children. This position was expressed more strongly in Camrose County than in Spirit River. In the case of Hillstream, whereas withdrawal of the Mennonite children was acceptable in principle as a religious or parental right, it was not acceptable in practice, as it threatened the ongoing public school system.

Members of school boards showed less ideological consistency. Persons interviewed were concerned about what was being learned in schools. They were particularly worried about the cost of schooling. In Spirit River, this concern was partially related to excessive school expenses, though the main focus was on government and lack of adequate government funding. It would seem that in Camrose this emphasis was reversed.

There were no significant differences in ideological orientation between the categories of respondents from the two communities. Ideology, as such, would therefore not seem to explain the differences in conflict in the two cases. What differences, then, exist at the structural level which might be considered causal of the different levels of conflict? There appear to be three of particular significance.

In the first place, the Hillstream school is much more seriously threatened by the withdrawal of Mennonite children. It has lost 25% of its school enrollment, while Georgia school has lost only 5%. Hillstream's teacher complement is reduced by two, a saving at the county level which Spirit River might envy (Spirit River was not able to make reductions in teaching staff), but which is certainly a threat at the local level. Whereas Hillstream may see a smaller and smaller school in the future, Georgia school continues to expand, with grade twelve having been added in 1977. The first case is one of retrenchment, the second of expansion.

It may be the case that this first factor, a structural factor associated strictly with schooling itself, would have been enough to generate the differences in levels of conflict. Yet further structural aspects must be noted. The Hillstream residents and district farmers see a threat not only to their school but to their whole community. A viable, active community of farmers, a relatively homogeneous community which lacked even the pluralism of separate schools, much less private schools, has had 25% of its population replaced by Holdeman Mennonite farmers. Unlike in the Spirit River district where Holdemans and others moved in at the same time, in Hillstream the Mennonites may be seen as "invaders" in the sense that term is used in ecological analysis of the Chicago "school of sociology."

Finally, in the Hillstream case the school teachers are an integral part of the community. As noted earlier, one is mayor of the local village, others own homes and farms in the area, and about half grew up in the locality. The community and the teachers can act relatively as one: the teachers' ideology may motivate their actions but their shared social relations with kin and neighbours partially enables them to transmit their ideology and to act as motivators in the community's organized resistance. By contrast, the teachers at Georgia are outsiders; living in mobile homes on the school grounds dramatically symbolizes this relationship. Some are beginning to farm in the area, and eventually school and community will no doubt become more articulated, but at present, neither the teachers' ideology of education nor their problems are shared through well-developed social relations.

There are, no doubt, other factors; other communities could be examined to determine if all of these elements are present. The role of the media in clarifying an ideological situation should be examined somewhat more closely, as well as the significance of an organized structure for parents to discuss their school concerns with school staff. However, the three components described above — structural threats to the school, a replacement threat to the community itself, and ideological conflict spread by one interest group to a larger community through shared social relations — are present in the situation where conflict is high and absent in the situation where conflict is low.

There are a myriad of other possible inferences and directions for further research suggested by this study. For example, it was stated earlier in the paper that "conflict around the school should not be conceptualized merely as a conflict of 'images' but rather as a conflict as to which of the other social institutions will, in fact, control schooling and thereby determine the nature of school" (Supra, p. 78). It is clear that the Holdemans have established their schools as extensions of the family and of religion, and that they very explicitly see control as appropriately and necessarily resting with these two institutions. Although they will accept guidance in relation to curriculum, they explicitly reject any assumption of the right of the state or state-certified teachers to exert control over these schools. Have they, then, created a school which in theoretical terms is a "type" of institution different from the public school? Would the literature on school effects,⁴ or consequence of schooling, apply to this type of school?

Professional teachers, on the other hand, do not see the appropriate authority for control of schooling to be the family or a religious institution. They see schooling control to be the province of those trained and licensed to perform that function. The degree to which teachers admit the legitimacy of the state to exert control is indeterminate: in terms of teacher certification, enforcement of legislation, or setting of official curriculum, teachers seem to accept state control because it is supportive of professional teacher interests and belief systems. As the Minister of Education's decision to accept a fourth category of approved private schools (one that does not require certified teachers) is not supportive of teacher interests, they are confused by it. In Rocher's terms, they are unable to give meaning to the Minister's actions which cannot be interpreted within the professional ideology.

Finally, one of the many omissions from the study is an analysis of the actions at the "centre," at the level of the government. By focusing research at the community level, we have no information on the beliefs or ideology and structural aspects influencing the decisions of the Attorney General's department, the

Minister of Education, the Cabinet, the Department of Education, or the Alberta Teachers' Association.

Notes

1. Rocher's (1972, p. 103) particular use of the term "ideology" needs amplification. He considers the term ideology to be close to W. I. Thomas' "definition of the situation," which is the way that members of a collectivity explain their present situation and give it meaning. However, he adds three additional points: (a) ideology has a systematic, coherent and organized form because it is explicit and verbalized; (b) ideology refers a great deal to values; (c) ideology has a conative function — it pushes or incites a collectivity to action, or at least directs this collectivity by providing goals and means.
2. The phrase "hegemony of beliefs" refers to an idea system which is so dominant or widespread that there is little or no conflict in a society. Thus, the alternative to Carlton's ideological conflict would be ideological hegemony.
3. We are indebted to our colleague, Professor Urion, for his observations on this point.
4. Alan Pomfret implicitly addresses a similar question in his 1977 paper on Newfoundland schooling. He basically argues that the structure of Newfoundland schooling reflects the "periphery" status of Newfoundland. This paper, partly a study of denominational schooling in Alberta, challenges some of his conclusions.

References

- Alberta, Bill of Rights, 1972.
- Archer, W. *Education and the Mennonites*. Unpublished paper, University of Alberta, February, 1978.
- Camrose Canadian*. Camrose, Alberta: February 1977 to February 1978.
- Carlton, R. A. Popular images of the school. *Alberta Journal of Educational Research*, 1974, 20, 59-74.
- Coulson, M., & Riddell, C. *Approaching sociology: A critical introduction*. London: Routledge and Kegan Paul, 1970.
- Epp, F. H. *Mennonites in Canada 1786-1920: The history of a separate people*. Toronto: Macmillan of Canada, 1974.
- Gamson, W. Rancorous conflict in community politics. *American Sociological Review*, 1966, 31 (1), 71-81.
- Gerth, H., & Wright Mills, C. *Character and social structure: The psychology of social institutions*. New York: Harcourt, Brace and World, Inc., 1953.
- Klassen, P. G. *A history of Mennonite education in Canada 1786-1960*. Thesis, University of Toronto, 1970.
- Lee, D., & Lapointe, J. Conflict over schools in a multi-ethnic society. In R. A. Carlton et al. (Eds.), *Education, change, and society*. Toronto, Gage, 1977.
- Pomfret, A. Historical change in a peripheral society: A working paper on Newfoundland. Paper presented at the meeting of the Canadian Sociology and Anthropology Association, Fredericton, New Brunswick, June, 1977.
- Rocher, G. *A general introduction to sociology: A theoretical perspective*. Toronto: Macmillan of Canada Ltd., 1972.

DAVID M. WILLIAMS

The University of British Columbia

A Study of Moral Education in Surrey, B.C. Secondary Schools

This paper reports the assumptions, procedures, findings and implications of the second year of a two-year study of moral education. Year 2 of the study involved 411 junior and senior secondary school students and 13 teachers. Groups in two experimental moral education conditions were compared with each other and with a control condition on a battery of measures of moral reasoning abilities and attitudes. The main findings were of significant treatment effects on normative reasoning and nonprejudice measures attributable to one of the moral education conditions. Hypotheses concerning treatment group differences in distinguishing factual and value claims and moral development were not confirmed. Implications of discrepant findings and procedural problems for further research and development are discussed. (Dr. Williams is an Associate Professor in the Social Studies Department, Faculty of Education, the University of British Columbia.)

Workers in the field of moral education and research face a range of problems, including (a) lack of agreement concerning the role of schools in promoting moral development, (b) lack of convincing theory concerning the requirements of mature moral judgment and action, (c) lack of effective means for the measurement and/or promotion of moral components, and (d) limited knowledge of the requirements for training or selecting teachers for work in moral education. Although problems of theory, test and curriculum development, and in-service education beset all areas of public education, they are particularly critical in a problematic and controversial field such as moral education.

The Association for Values Education and Research (AVER, Faculty of Education, U.B.C.) has conducted a multidimensional research and development project in cooperation with junior and senior secondary schools in Surrey, B.C. The project has been sufficiently comprehensive and flexible to allow for experimental comparisons of alternative approaches to moral development, as well as concurrent development work in curriculum, measurement, and teacher training.

The activities, problems, and findings of Year 1 were reported in the 1975 Yearbook of The Canadian Society for the Study of Education (Williams, 1975).

The purpose of the present paper is to describe the assumptions, activities, findings and implications of Year 2 of the AVER-Surrey Project.

Assumptions Underlying the AVER-Surrey Project

AVER's work is based on the assumption that any adequate account of the requirements of mature moral judgment and action must be at least as comprehensive as that shown in Figure 1. Those of you who are familiar with the moral education literature will recognize the influence of John Wilson (1972), Coombs and Meux (1971), and earlier AVER research (AVER, 1974, 1975) on this "working list" of moral components. AVER assumes, in addition, that there are developmental constraints on the acquisition of moral reason-giving (Kohlberg, 1971), social perspective or role-taking (Selman, 1975), and other moral components.

-
1. Disposition to regard others as equals and to give weight to one's own and others' interests (CONCERN FOR OTHER PEOPLE).
 2. Ability to know what others (and yourself) are feeling and what their (and your) interests are (EMPATHY/ROLE-TAKING).
 3. Knowledge of facts relevant to moral choices (SOCIAL KNOWLEDGE).
 4. The 'know-how' to act effectively in social contexts (SOCIAL SKILLS).
 5. Ability and disposition to give reasons for moral judgments and actions *and* to assess reasons given in terms of factual warrant *and* moral justification (REASON-GIVING, FACT ASSEMBLY and PRINCIPLE TESTING).
 6. Sensitivity to morally hazardous situations, including
 - a. actions that may have consequences for others that one could not accept for oneself, and
 - b. actions that may have disastrous consequences if everyone were to engage in them (MORAL POINT-OF-VIEW).
 7. Resolution to act on justifiable moral judgments: to do what one has decided is right and to refrain from doing what one has decided is wrong.
-

Figure 1. Moral abilities, attainments, and dispositions.

These assumptions suggest important implications for moral education theory and research. They imply, for example, that it is unlikely that a single instructional strategy will promote all of the moral components. Another implication is that developmental constraints will limit the degree to which some, if not all, of the moral components will be amenable to instructional manipulation within limited treatment periods. Finally, these assumptions imply for AVER that the structural analyses of global stages being undertaken by researchers in the developmental psychological tradition, such as Kohlberg and Selman, can at best provide only a partial picture of moral development and education. The ability of researchers to weigh the relative importance of such components as those identified in Figure 1 in predicting mature moral judgment and action will depend in part upon advances in the measurement and promotion of discrete components. This moral components approach, rather than the developmental psychological orientation, is predominant in the AVER-Surrey Project.

Curriculum Development: Experimental Conditions

In Year 1 of the project, AVER developed curricula, trained teachers and evaluated the effects of three approaches to values and moral education with respect to each other and to a non-values, control curriculum. Junior secondary school students in grades eight, nine and ten were the subjects of the Year 1 study. As a result of Year 1 activities and findings, a second study involving further curriculum development, teacher training, test modification and comparison of approaches was undertaken in Year 2 of the project.

On the basis of Year 1 findings, it was decided to combine elements of the Meux (1974) programmed text materials (known in Year 1 as the Meux condition, or approach) with the reasoning elements of the AVER condition to comprise a modified AVER condition for implementation and evaluation in Year 2. The Kohlberg condition was retained in modified form in view of Year 1 and independent evidence (Kohlberg, 1971) suggesting the promise of this approach for the enhancement of some of the moral components. Unlike Year 2, funding did not permit curriculum development and teacher training for a non-values control condition. However, it *was* possible in one of the participating Year 2 schools to assign a proportion of grade 11 and 12 students to regular English and Social Studies classes, which comprised a quasi-control condition. In effect, therefore, Year 2 involved curriculum development and teacher training for two experimental conditions, AVER and Kohlberg. Testing was conducted with students under three conditions, AVER, Kohlberg, and control.

Experimental Condition 1. The AVER Approach

AVER believes that a morally autonomous, or morally educated, individual must bring to bear upon any moral problem a number of attainments, abilities and

Theme	Tasks
Unit 1: Status of Woman	Facts/Values Distinction; Clarifying Value Objects; Role Playing; Practical Syllogisms.
Unit 2: Population	Fact/Value Distinction, Identifying Point of View; Fact Assembly: Rudimentary Procedure.
Unit 3: Pollution	Fact/Value Distinction; Role Playing; Fact Assembly: Rudimentary Procedure; Principle Testing: Role Exchange Test.
Unit 4: Prejudice	Determining Relevant Differences; Fact Assembly Chart; Practical Syllogism; Principle Testing; Role Exchange Test; Universal Consequences Test; New Cases Test; Subsumption Test.
Unit 5: War and Revolution	Clarifying Value Objects; Point of View: Moral Viewpoint; Practical Syllogisms; Indicators of Relevance.
Unit 6: Old Age	Fact/Values Distinction; Point of View; Role Playing; Fact Assembly Chart; Practical Syllogisms; Principle Testing: Role Exchange Test; New Cases Test and Subsumption Test.
Unit 7: Capital Punishment	This was designed as the final unit in which all AVER Tasks were reviewed.

Figure 2. AVER tasks by theme and unit.

dispositions. Various exercises designed to foster the acquisition of moral reasoning and role-taking components were incorporated into the AVER condition curriculum materials. AVER exercises included: (a) distinguishing factual and value claims, (b) identifying point-of-view, (c) fact assembly and rudimentary procedure, (d) the practical syllogism, (e) principle testing, and (f) role playing.

AVER units focusing on seven moral value issues were assembled. The ideal of having each unit focus upon a particular ability as well as serve as a cumulative review of the other moral components was only partly achieved, because of limited time and resources. Figure 2 identifies the tasks emphasized in individual AVER units.

Experimental Condition 2. The Kohlberg Approach

Lawrence Kohlberg (1971) claims on the basis of extensive longitudinal, cross-cultural and experimental research that moral reasoning develops through an invariant sequence of stages. He and Turiel (1973) have found that students can comprehend and tend to prefer reasoning at one stage higher than their own dominant stage. Furthermore, Kohlberg claims that movement towards the next

Level 1 — Premoral

- Stage 1 — Orientation toward punishment and unquestioning deference to superior power. The physical consequences of action regardless of the human meaning or value determine its goodness or badness.
- Stage 2 — Right action consists of that which instrumentally satisfies one's own needs and occasionally the needs of others. Human relations are viewed in terms like those of the market-place. Reciprocity is a matter of "you scratch my back and I'll scratch yours", not of loyalty, gratitude or justice.

Level II — Conventional Role Conformity

- Stage 3 — Good-boy-good-girl orientation. Good behaviour is that which pleases or helps others and is approved by them. There is much conformity to stereotypical images of what is majority or "natural" behaviour. One seeks approval by being "nice".
- Stage 4 — Orientation toward authority, fixed rules a of the social order. Right behaviour consists of doing one's duty, showing respect for authority and maintaining the given social order for its own sake. One earns respect for performing dutifully.

Level III — Postconventional Morality

- Stage 5 — Social-contract orientation, generally with legalistic and utilitarian overtones. Right action tends to be defined in terms of general rights and in terms of standards which have been critically examined and agreed upon by the whole society. There is a clear awareness of the relativism of personal values and opinions and corresponding emphasis upon procedural rules for reaching consensus.
- Stage 6 — Orientation toward the decisions of conscience and toward self-chosen ethical principles appealing to logical comprehensiveness, universality and consistency. They are not concrete moral rules like the Ten Commandments. Instead, they are universal principles of justice, of the reciprocity and equality of human rights, and of respect for the dignity of human beings as individual persons.

Figure 3. Kohlberg's stages of moral development.

higher stage of moral reasoning can be enhanced if subjects are given opportunities to hear “+1” reasoning during group discussions of difficult moral choices, or dilemmas. A brief description of Kohlberg’s six stages of moral development is given in Figure 3.

Curriculum materials for the Kohlberg condition focused upon the same unit themes as the AVER condition. However, the materials comprised Kohlbergian moral dilemmas relating to each of the unit themes. These were presented in print, sound filmstrip, and role playing formats. Questions accompanying the dilemmas encouraged pupils to make choices and to give reasons for those choices. Teachers were expected to promote reason-giving at a variety of stages of Kohlberg *and* to promote lively exchange of reason-giving on the part of pupils.

Teacher Training

Although a total of 13 teachers participated in Year 2 of the AVER — Surrey Project, only the 11 teachers assigned to one or other of the experimental conditions took part in training sessions conducted prior to the implementation period. The two control condition teachers were not trained by AVER personnel. Control teachers implemented their regular English or Social Studies programmes. It must be acknowledged, therefore, that unlike Year 1, comparisons of the experimental conditions with the control condition fail to control for a possible “Hawthorne Effect”¹ by giving equal attention to all teachers and subjects.

Teacher participants assigned to the AVER or Kohlberg conditions worked with project personnel in up to eight day-long training sessions. Teachers assigned to each condition undertook differentiated training specific to the condition to which they had been assigned.

AVER Condition Training Sessions

Five teachers met with project personnel for training sessions. Programmed text materials on evaluative reasoning (Meux, 1974) were utilized in the early sessions to elucidate the nature of value judgments, types of value judgments, moral and non-moral points of view, and the use of rudimentary procedure in the assembly of facts relevant to value judgments.

Teachers viewed a videotape of an AVER teacher demonstrating the use of syllogistic reasoning in classroom discourse. Teachers undertook similar tasks to those they would later administer to their students. Some emphasis was placed on the implementation of role-taking tasks in the classroom. Trainees were also instructed in the utilization of four methods of testing principles implied by moral judgments, the role exchange test, the universal consequences test, the new cases test, and the subsumption test.

Later training sessions involved examination and discussion of curriculum materials to be utilized in implementation of the AVER condition.

Kohlberg Condition Training Sessions

Six teachers took part in Kohlberg training sessions. These sessions were of roughly the same duration and extended over the same months prior to the implementation period as the AVER condition training sessions.

Numerous tasks were undertaken by trainees under the direction of project personnel to establish a working knowledge of Kohlberg’s stages of moral

reasoning. Kohlbergian dilemmas were discussed with examples given of typical responses at each stage. Teachers were encouraged to compose their own dilemmas and to present appropriate arguments for choices at each of the moral reasoning stages.

Once the teachers were familiar with the basic stages of Kohlberg's theory of moral reasoning development, they were instructed in various methods of presenting dilemmas to groups of students. This instruction included the viewing of videotaped demonstrations in Kohlbergian moral education conducted by leaders experienced in dilemma discussion techniques. Later training sessions were devoted to the presentation of draft units of the Year 2 Kohlberg condition. Teachers and project personnel discussed potential implementation problems and suggestions for improvement of the curriculum materials.

Evaluation Design and Subjects

Table 1 shows the approximate number of students by treatment condition and school for whom complete pretest-posttest data were available on one or more of the measures. Differences in cell sizes within the schools and terms of the school year are related to the need to assign subjects to treatment groups of approximate "class" size within timetable blocks. Although it was possible to assign subjects randomly to treatment conditions within timetable blocks, it was not possible to assign them randomly across timetable blocks. The availability of project teachers within timetable blocks necessitated the unequal distribution of subjects to treatment groups shown in Table 1. The junior secondary design is incomplete in the sense that no non-values, control groups were available for comparison with the experimental conditions. Nevertheless, it will be noted that for all hypotheses the one experimental group served as control for the other. In the senior secondary school design, the possibility of comparing experimental groups with non-experimental control groups was also available.

TABLE 1
NUMBER OF SUBJECTS BY TREATMENT CONDITION, GRADE, AND SCHOOL

School/Grade	AVER	Kohlberg	Control
Junior Secondary School: 3rd Quarter (Grade 8)	52	27	--
Junior Secondary School: 4th Quarter (Grade 8)	27	52	--
Senior Secondary School: 2nd Semester (Grades 11/12)	66	88	69

Table 2 shows differences in testing, treatment period, and intensity of treatment by school and school term. These differences persuaded the researchers to analyse separately data gathered in different terms and schools. In effect, then, this report describes findings of three discrete studies, junior secondary third quarter, junior secondary fourth quarter, and senior secondary, relevant to evaluation of the effects of the experimental conditions. It will be noted that the

TABLE 2
TREATMENT PERIOD, TESTING AND INTENSITY OF TREATMENT BY SCHOOL

School/Term	Pretest	Treatment Period	Posttest	Intensity
Junior Secondary 3rd Quarter	February Week 1	February–March Inclusive	April Week 2	Concentrated (1 hour per day)
Junior Secondary 4th Quarter	April Week 1	April–May Inclusive	June Week 2	Concentrated (1 hour per day)
Senior Secondary 2nd Semester	February Week 1	February–May Inclusive	June Week 1	Distributed (50% experimental; 50% regular curriculum; 6 hours per week)

overall design does not permit the partialling out of effects attributable to concentrated versus distributed implementation of the treatment conditions.

Batteries of tests were administered during pretest and posttest periods. Monitoring of treatment implementation was done by means of direct observation by project personnel utilizing an observation checklist (AVER, 1975).

Test Development

A major task of Year 2 was the locating, modifying, and further development of measures of the various dispositions and abilities associated with moral competence.

In Year 2 of the project, the following measures were utilized in the pre and/or posttest batteries of the AVER-Surrey Project: the Defining Issues Test (Rest, 1974; modified by AVER, 1975), the Facts/Values Test (AVER, 1974; modified by AVER, 1975), the Normative Reasoning Test (AVER, 1974; modified by AVER, 1975), a Prejudice Test (Adams, 1975, Forms A and B), and the Social Desirability Scale (Crandall, Crandall & Katkovsky, 1965).

The Defining Issues Test (DIT) is purported to measure moral development. Responses are keyed to Kohlberg’s stage types. The modified version of the DIT utilized by the AVER-Surrey Project differentiates between conventional and principled moral reasoning on five dilemma stories, including the Heinz, Escaped Prisoner, Newspaper, Webster, and Doctor’s Dilemmas. Rest claims that “P”², or principled morality, scores based on the first three of these dilemmas correlate .93 with P scores based on the six dilemma long-form of the DIT. Evidence concerning the reliability and validity of the DIT is provided in Rest’s *Manual for the Defining Issues Test* (1974). It is worth noting that AVER modified the DIT by adding two dilemmas accompanied by response items judged to be more sensitive to lower stage (i.e., pre-conventional) reasoning than the Rest dilemmas and response items. However, P scores reported here are based only on the dilemmas and items developed by Rest.

The Normative Reasoning Test (AVER, 1975) utilized in Year 2 of the project comprised 30 items in which relationships between premises and conclusions in normative arguments were sound or unsound. Subtest A comprised 20 multiple-choice items for which pupils were required to select the soundest conclusions following from a value premise and a factual premise. Hoyt (1941)

estimates of reliability for samples of junior and senior secondary subjects in Surrey, B.C. schools range from .71 to .78 for subtest A. Subtest B comprised 10 multiple-choice items for which pupils were required to select the reason or premise that best completed a normative argument. Hoyt estimates of reliability for Subtest B range from .61 to .81. Overall estimates of reliability for the total 30-item test range from .78 to .86, according to Hoyt estimates of reliability, and .68 to .76, according to Cronbach's alpha (1951). For purposes of the group comparisons made in this study, such reliability coefficients were judged acceptable.

The Facts and Values Test (AVER, 1975) is a modification of the test used in Year 1. In pretest form, it comprised 20 items, but in posttest form it comprised 30 items that subjects were required to identify as either factual or value claims. Hoyt estimates of reliability range from .69 to .79 for four samples of junior and senior secondary school subjects in Surrey schools.

The Prejudice Test (Adams, 1975) utilized in the posttest battery is a 27-item Likert-type measure of subject's prejudicial attitude towards an ethnic group to whom considerable community hostility was being directed at the time of the study. Hoyt estimates of internal consistency ranging from .94 to .95 were obtained from samples of Surrey, B.C. junior and senior secondary school subjects on this measure.

The Social Desirability Scale (Crandall, Crandall & Katkovsky, 1965) is a 26-item measure of a subject's disposition to deceive others about oneself. As such, it is a measure of frankness or honesty in self-reports. A high score indicates a high disposition to deceive. Hoyt estimates of reliability for the measure range from .76 to .80 for samples of Surrey, B.C. junior and senior secondary school subjects.

Year 2: Hypotheses and Findings

The experimental/evaluation portion of Year 2 of the AVER-Surrey Project was designed to test certain hypotheses and to check for unintended effects of the experimental conditions on prejudice and disposition to deceive. Kohlbergian theory and Year 1 findings suggested the following hypotheses with respect to treatment group differences:

1. On the Normative Reasoning Test, it was predicted that
AVER > Kohlberg = Control.
2. On the Facts and Values Test, it was predicted that
AVER > Kohlberg = Control.
3. On the Defining Issues Test, it was predicted that
Kohlberg > AVER = Control.

Recent literature on curriculum evaluation (Scriven, 1967; Worthen & Saunders, 1973) suggests that data should be gathered concerning unintended and undesirable effects of programmes. In the case of the AVER and Kohlberg conditions, possible negative effects of classroom discourse about prejudice were of concern. Hartshorne and May (1929-30) and others have reported evidence suggesting that some character training/moral education programmes increase rather than reduce pupil tendencies to deceive teachers and evaluators about their sentiments and conduct. In order to assess unintended treatment effects on these variables, the Prejudice Test (Adams, 1975) and Social Desirability Scale (Crandall, Crandall & Katkovsky, 1975) were administered to subjects. Results of

analyses of variance of treatment effects among senior secondary school subjects (the only portion of the design permitting experimental-control group comparisons) are reported here.

Treatment Group Differences on the Normative Reasoning Test

On the Normative Reasoning Test, it was predicted that AVER > Kohlberg = Control. Table 3 shows treatment group differences and an analysis of covariance of treatment effects on senior secondary school students. The finding of a significant treatment effect and of significant differences between the AVER and Kohlberg subjects and the AVER and control subjects, as well as the direction of those differences, tends to confirm hypothesis 1 for senior secondary students.

TABLE 3
ANALYSIS OF COVARIANCE AND TREATMENT GROUP MEANS AND STANDARD DEVIATIONS ON NRT POSTTEST WITH NRT PRETEST SCORES USED AS COVARIATES: SENIOR SECONDARY SCHOOL SUBJECTS

Source	D.F.	M.S.	F	p
Treatment	2	214.28	9.35	.0001*
Error	199	22.92		
	Unadjusted $\bar{X}$	Adjusted $\bar{X}$	S.D.	N
AVER	23.95	23.13	] 3.66	61
Kohlberg	19.80	19.85		88
Control	19.24	20.08		54

] Significant ($\alpha = .05$) treatment group differences according to Newman-Keuls Multiple Range Test.

* Significant treatment effect.

Analyses of treatment effects of the Normative Reasoning posttest for junior secondary students in the 3rd and 4th quarters of the school year were confounded by significant treatment group differences on the pretest, despite random assignments of subjects to the AVER and Kohlberg conditions. When differences on the pretest were controlled by means of analyses of covariance, results revealed no significant treatment group differences among junior secondary school comparison groups.

Treatment Group Differences on the Facts and Values Test

On the Facts and Values Test it was predicted that AVER > Kohlberg = Control. However, analyses of covariance of posttest scores with pretest scores used as covariates revealed findings contrary to this prediction in one of the three populations sampled.

Table 4 shows data relevant to hypothesis 2 from junior secondary students in the fourth quarter of the school year. The difference, favouring as it does the Kohlberg condition, is contrary to hypothesis 2.

TABLE 4
ANALYSIS OF COVARIANCE AND TREATMENT GROUP MEANS AND STANDARD
DEVIATIONS ON THE FVT POSTTEST WITH THE FVT PRETEST SCORES
USED AS COVARIATES: JUNIOR SECONDARY SUBJECTS, 4th QUARTER

Source	D.F.	M.S.	F	p
Treatment	1	153.89	7.49	.008*
Error	77	20.55		

	Unadjusted $\bar{X}$	Adjusted $\bar{X}$	S.D.	N
AVER	9.58	9.77	3.67	26
Kohlberg	12.83	12.74	5.13	54

* Significant treatment effect.

Treatment Group Differences on the DIT

Analyses of covariance for the senior secondary and two junior secondary samples failed to confirm the prediction of significant treatment group differences in favour of the Kohlberg condition. In fact, with all three samples of subjects, the direction of mean differences favoured the AVER over the Kohlberg condition, although these may very well have been chance differences.

Treatment Group Differences on the Prejudice Test (Adams, 1975)

The prejudice unit comprised only a small portion of the experimental conditions in Year 2 of the AVER-Surrey Project. It is not unreasonable, however, to assume that moral education programmes dealing with the issue of prejudice might have some effect, either positive or negative, on prejudicial attitudes towards minority groups. In the interests of evaluating possible experimental-control group differences on this variable, a prejudice test was administered as part of the posttest battery to senior secondary school students.

High scores on the Prejudice Test indicate greater prejudice than do low scores. Table 5 shows a significant treatment effect, with the direction of mean differences suggesting that students in the control and Kohlberg conditions were more prejudiced than the AVER students. However, the Newman-Keuls test failed to confirm the significance of particular comparisons among the treatment conditions.

Treatment Group Differences on the Social Desirability Scale

The Social Desirability Scale (Crandall, Crandall & Katkovsky, 1965) can be interpreted as a measure of subjects' disposition to deceive others about themselves. As such, it taps a construct of great potential interest to researchers in moral education. Analysis of variance of treatment group differences in social desirability scores reveals no significant differences among comparison groups. This finding increases confidence that student self-reports on the Prejudice and Normative Reasoning Tests were not differentially affected by disposition to deceive.

TABLE 5
ANALYSIS OF VARIANCE AND TREATMENT GROUP MEANS AND
STANDARD DEVIATIONS ON THE PREJUDICE TEST:
SENIOR SECONDARY STUDENTS

Source	D.F.	M.S.	F	p
Treatment	2	2686.83	3.97	.02*
Error	235	676.22		

	$\bar{X}$	S.D.	N
AVER	73.54	24.59	71
Kohlberg	81.41	25.09	91
Control	85.45	28.27	76

* Significant treatment effect.

Monitoring of Conditions During Treatment Period

A total of 54 observations of class sessions were made utilizing an observation checklist designed by project personnel (AVER, 1975). No estimate of interobserver reliability in the use of the instrument is presently available.

Although no tests of significance were applied to observation checklist data, some differences reported by project observers were consistent with those expected between AVER, Kohlberg, and control conditions. Descriptive analyses show that Kohlberg teachers posed dilemmas, encouraged reason-giving at a variety of stages, rephrased responses to clarify stage level, and refused to let pupils “escape the dilemma” 50% more frequently as a proportion of their teaching behaviours than did AVER teachers. The proportion of teaching behaviours in the facts/values distinction, syllogistic reasoning, and principle testing categories was 50% higher for AVER than for Kohlberg teachers. In the use of other categories of behaviour specified by the observation checklist, experimental and control teachers were roughly similar.

Conclusions and Implications

A number of gains in research and practice in moral education have resulted or are likely to result from the findings and activities of the AVER-Surrey Project.

Moral education curriculum materials focusing on such issues as population, pollution, the aged, capital punishment and prejudice have been developed and evaluated. These materials are now being revised and edited in preparation for publication.

Teacher training sessions and classroom observations of teachers working with pupils provide, at least, pilot data concerning the effectiveness of our in-service training methods. Year 1 and 2 observations suggest that some teacher training methods were effective, whereas others were not. Differences among teachers trained in the same manner may be related to main effects and/or interactions of teacher characteristics, experimental condition, pupil characteristics, and teacher training methods. We cannot discount the possibility that not all teachers may be

suited to the teaching of certain types of programmes in values and moral education. Research into the effectiveness of teacher training methods in moral education and into differences among teachers in the various moral components is indicated.

The needs of research and theory in moral education have been advanced by the direct comparison of differentiated approaches to values and moral education *and* by the development, modification, and identification of tests suitable to assessing some of the effects of programmes in moral values education. Over the two years of the project, normative reasoning and facts-values tests possessing a moderately high degree of internal consistency have been developed. Tests of dispositions relevant to particular moral education programmes (i.e., the Prejudice Test and the Social Desirability Scale) have been identified, modified and utilized. The project has provided new data concerning the internal consistency of these measures. Although attempts to develop a valid and reliable measure of role-taking and to make the DIT sensitive to moral reasoning development from the preconventional to the conventional level have not yet reached fruition, considerable data have been gathered that may yet lead to advances in these areas of measurement.

The expected superiority of the AVER condition in promoting the ability to recognize sound and unsound normative arguments was partially confirmed. Among senior secondary school students, the AVER treatment was significantly more effective than the Kohlberg and control treatments in promoting this ability. It may be that the failure of the AVER condition to yield a significant treatment effect on the Normative Reasoning Test among junior secondary school students is related to developmental constraints on the acquisition of the skill being tested. However, it is also possible that chance or tester-induced systematic differences existing prior to the treatment period confounded the discovery of significant treatment effects.

Although not predicted, the finding of a significant AVER effect on the Prejudice Test is encouraging. It is possible that the AVER condition was more effective than the Kohlberg condition in promoting dissonance concerning subjects' prejudicial beliefs. However, such speculation must await further investigation, including replication of present findings.

Perhaps the most inexplicable finding of Year 2 is that of a significant treatment effect among fourth-quarter junior secondary students on the Facts and Values Test *favouring the Kohlberg over the AVER condition*. In Year 1, the Meux condition proved to have a significant effect on promotion of the facts-values distinction. On the strength of this finding, elements of the Meux programmed text materials were incorporated into the AVER condition for Year 2. Certainly the AVER treatment was the only Year 2 condition explicitly designed to enhance the skill of differentiating factual and value claims. However, it appears that the AVER condition might actually have depressed scores on the modified Facts and Values Test in Year 2 (!). A possible explanation is that the AVER program made subjects sufficiently sensitive to different classes of "factual" and value claims that the task presented in the test became more, rather than less, difficult to AVER subjects. However, this is merely conjecture. Certainly, the discrepant Year 1 and 2 findings concerning treatment effects on the Facts and Values Test warrant further investigation.

The expected superiority of the Kohlberg condition in promoting moral

reasoning development (as measured by the DIT) was *not* confirmed. The failure to yield significant results may be attributable to such factors as lack of power in the treatment, the relative insensitivity of the DIT to stage transition from the premoral to conventional level of moral reasoning or, even, extreme differences of cell sizes among the treatment groups analyzed.

The general lack of effectiveness of the Kohlberg condition and of the AVER condition among junior secondary school students requires further investigation. Did the insensitivity of the DIT as a measure of moral reasoning development at lower stages of Kohlberg bias the Year 2 results against the Kohlberg condition and/or was the form of the Kohlberg condition utilized in this study lacking in power? The ineffectiveness of the AVER condition on junior secondary students may well be related to the level of sophistication and/or difficulty of AVER materials and activities. The need to design AVER activities and materials specifically adapted to the reading and maturity level of younger students is certainly indicated by Year 1 and Year 2 findings.

This paper reports work conducted by the Association for Values Education and Research (AVER, Faculty of Education, the University of British Columbia) in a project funded by the Vancouver Foundation, the Educational Research Institute of B.C., the Surrey School District, and the Office of Field Development of the University of British Columbia Faculty of Education. The author presents this report on behalf of his colleagues and co-workers in AVER.

Notes

1. "Hawthorne Effect" may be defined as improved performance associated with participation of Subjects in situations involving the special attention accorded to experimental treatment groups (Sax, 1968, 339-340).
2. Rest claims that the P score, an index derived from responses to DIT items, can be interpreted as "the relative importance attributed to principled considerations" (Rest, 1974, 2-2).

References

- Adams, D. K. Prejudice test. Unpublished Likert-type attitude scale. University of British Columbia, Faculty of Education, 1975.
- AVER. *Report No. 5: Interim report of an experiment in moral education at Surrey, B.C.* Vancouver: Association for Values Education and Research, 1974.
- AVER. *Report No. 6: Final report of a study in moral education at Surrey, B.C.* Vancouver: Association for Values Education and Research, 1975.
- Coombs, J. R., & Meux, M. Teaching strategies for value analysis. In L. E. Metcalf (Ed.) *Values education: Rationale, strategies and procedures*. Washington, D.C.: National Council for the Social Studies, 1971.
- Crandall, B. C., Crandall, V. J., & Katkovsky, W. A children's social desirability questionnaire. *Journal of Consulting Psychology*, 1965, 29, 27-36.
- Cronbach, L. J. Coefficient alpha and the internal structure of tests. *Psychometrika*, 1951, 16, 297-334.
- Guidance Associates. *First things: Values*. Pleasantville, N.J.: Guidance Associates, 1972. (6 sound filmstrips and discussion guides.)
- Hartshorne, H., & May, M. A. *Studies in deceit* (Vol. I). New York: Macmillan, 1929-30.
- Hoyt, C. Test reliability estimated by analysis of variance. *Psychometrika*, 1941, 6, 153-160.
- Kohlberg, L. Stages of moral development as a basis for moral education. In C. M. Beck, B. S. Crittenden & E. M. Sullivan (Eds.) *Moral education: Interdisciplinary approaches*. New York: Newman Press, 1971.
- Metcalf, L. E. (Ed.) *Values education: Rationale, strategies and procedures*. Washington, D.C.: National Council for the Social Studies, 1971.
- Meux, M. & Associates. *Rational value decisions and value conflict resolution: A handbook for teachers*. Salt Lake City: Granite School District, 1974.

- Rest, J. *Manual for the defining issues test*. Minneapolis: University of Minnesota, 1974.
- Sax, G. *Empirical foundations of educational research*. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1968.
- Scriven, M. The methodology of evaluation. In R. Tyler, R. Gagne, & M. Scriven (Eds.) *Perspectives of curriculum evaluation*. A.E.R.A. Monograph 1. Chicago: Rand McNally, 1967.
- Selman, R. L. Level of social perspective taking and the development of empathy in children: Speculation from a social-cognitive viewpoint. *Journal of Moral Education*, 1975, 5 (1), 35-43.
- Turiel, E. Stage transition in moral development. In R. M. W. Travers (Ed.), *Second handbook of research on teaching*. Chicago: Rand McNally, 1973.
- Williams, D. AVER in Surrey: An approach to research and development in moral education. *C.S.S.E. Yearbook*, 1975, 2, 62-75.
- Wilson, J. *Practical methods of moral education*. London: Heinemann, 1972.
- Worthen, B. R., & Sanders, J. R. *Educational evaluation: Theory and practice*. Worthington, Ohio: Charles A. Jones Publishing, 1973.

EVA SHAW

The University of Calgary

The Effects of Attentional Focus on Graphic Discrimination

This study compared the effectiveness of four methods of training young children in the skills of visual discrimination necessary to the acquisition of letter knowledge. Subjects were all children between the ages of four years and four years six months, attending day care centres subsidized by the City of Calgary, Alberta. Data were analyzed using a multivariate analysis of variance, with a probability of .05 as the level of significance. Within the limitations of the study, results indicated that training in oral description of critical cues of letters contributed to superior achievement in learning to match letters. In light of the findings of the study, an amendment to Gough's model of reading (Gough, 1972) was suggested. (Dr. Shaw is in the Department of Educational Curriculum and Instruction, Faculty of Education at the University of Calgary.)

The transition from the understanding of language through audition to the understanding of language through vision is a complex process in which the eye must gradually supersede the ear as the organ of reception. In order to learn to read, it is essential that the learner be able to discriminate letters. Internalization of letters follows the hierarchical stages of matching (awareness of the sameness of two letters), identification (the ability to single out a specified letter from a group of letters on request), and recognition (the ability to write and name a letter on request from memory). In the process of learning to match letter forms, young children are asked to compare symbols which probably bear no associational meaning, and certainly bear no common associational meaning. Thus, the task is strictly perceptual. The child needs to be aware of line, slope, curve, and direction. He needs also to understand clearly the purpose of the task required of him. He cannot match unless he has a true perception of "same" and "different." As Gibson, Gibson, Pick and Osser (1962) pointed out, at no previous time in a child's life has the directionality of an object affected its label.

Recent research has shown an increased interest in the desire to describe clearly the process of interactions through which the reader absorbs information from print. Many of the subprocesses of reading appear to take place simultaneously and with greater or lesser effect on the end-product, according to the proficiency of the

reader. Studies of the complexity of the reading process have resulted in the construction of various models of reading. While there are differences among the models, most of them use single word recognition as the starting point of the reading process. Mackworth (1972), however, stated that if the letter learning itself has been inadequate, there is a strong possibility that the learning of simple word pattern and the linkage of these patterns with their sounds will present serious difficulties to the beginning reader.

The questions then arise as to how children begin the task of letter recognition. What, if any, scanning devices do they employ when making visual comparisons? What particular information do they select and retain from a symbol that may be totally devoid of meaning for them?

It is possible that some order may be brought into the child's perceptual environment through the use of language. There is general agreement that at some time in a child's development, the power of thought exceeds the power of language. Since language is the verbal expression of thought, it is obvious that what a child says, he is able to think. What is questionable is whether a young child's verbal capacity can be increased to express the complexity of his thoughts. Language was defined by Tansley (1970) as a "regulator of behaviour" and thus he saw perception, sharpened by language, leading to more effective conceptualization. The use of language leading to the verbal expression of strategies of comparison could provide the child with a means of reducing doubts in a manner which might reinforce his discriminatory skills.

Purpose

Research evidence supports the importance of visual discrimination training in beginning reading. The main purpose of this study was to investigate the processes by which a young child discriminates visually, since this is the first step towards letter recognition, and to investigate the effect of a verbal response on visual discrimination. An attempt was made to gain knowledge of the invariant features by which prereaders distinguish letter patterns.

Background to the Problem

Among recent models of the reading process the one by Gough (1972) presents a point of view emphasizing the need to investigate the process of letter recognition in the early prereading stages (Fig. 1).

The starting point of Gough's model is the eye fixation. Each fixation lasts approximately 250 milliseconds before the next saccade is made. As each fixation is achieved, the visual pattern reflected on the retina activates the visual system and results in the formation of an icon. Gough saw the icon as "an 'unidentified' or 'precategorical' visual image" (Gough, 1972, p. 332) capable of including everything in an oval shape roughly two inches wide and one inch high. Reputable research findings concerning the existence and the function of the icon were accepted by Gough, but he drew attention to the minimum of knowledge on the formation of the icon. Available data led Gough to assume that the initial fixation results in an icon which, for practical purposes, would include 15 to 20 letters and spaces of a sentence. Recognition of the icon would be complete in about 100 milliseconds and decay of the icon would occur in about 250 milliseconds as it becomes replaced by the new icon resulting from the next fixation. As this activity is taking place,

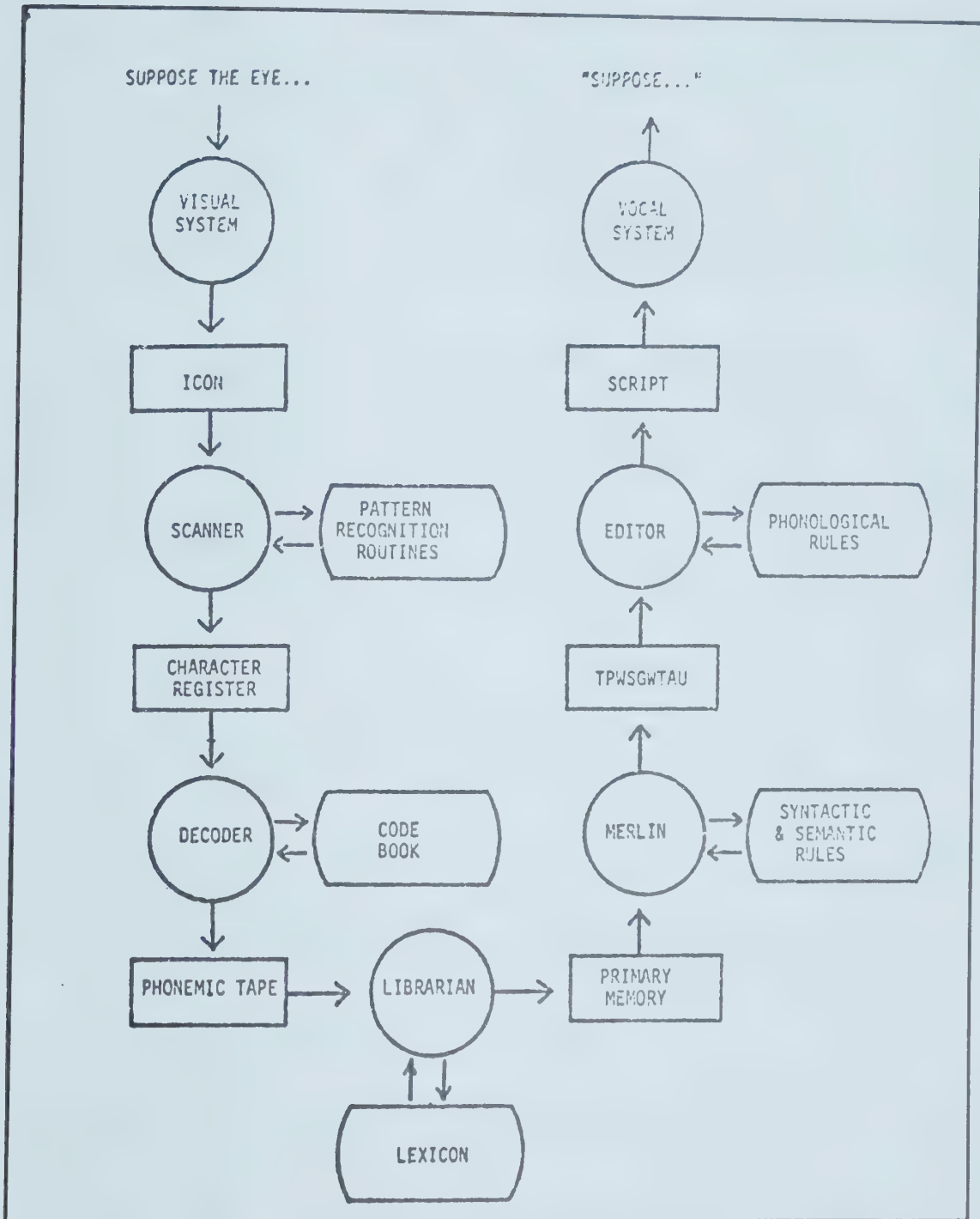


Figure 1. A model of reading. (From Philip B. Gough, "One Second of Reading" in J. F. Kavanagh and I. G. Mattingley [Eds.] *Language by Ear and by Eye: The Relationships Between Speech and Reading*. Cambridge, Mass.: M.I.T. Press, 1972. Reprinted by kind permission of the author and publishers.)

familiarity of the reader with the precategorical image results in identification. Gough assumed that the shapes are identified as letters.

When referring specifically to the prereader, Gough described him as being already equipped with several of the components needed for reading: a visual system which produces an icon, a vocabulary, a degree of comprehension, and a phonological system. However, what the child lacks is a scanner, a character

recognition device, and a decoder. Gough saw the decoder as a mechanism by which meaning becomes attached to the character.

Gough discussed the problem of character recognition and asserted that it was a problem, for while it had been shown that young children can discriminate letters, it was also obvious that children confuse letter shapes in successive presentations as they normally occur in the reading situation. There is no real information at this time about the features by which young children differentiate letters. Gough stated that if the subcomponents of the character recognition device could be identified, it might be possible to find out which subcomponents are present in the beginning reader.

The character register (Gough, 1972, p. 335) to which Gough referred appears to be a system of storing significant signals received serially from the icon, every ten or twenty milliseconds, in the form of letters. Signals, however, may emerge as letters only if the reader is able to label the signals as such. The fact that the prereader may have had no previous perceptual experience with the symbol confronting him makes it difficult for him to use retrieval mechanisms. Gough saw these mechanisms as the connecting link between the character register and the representations from print that are comparable with the reader's speech comprehension.

Since we do not know in what form the printed symbol, *per se*, appears in the icon, it is important to know how the critical features of letters are identified by a child. Gough appeared reluctant to categorize reversal and rotation errors as symptoms of reading disability; he accepted that they are not exclusive to poor readers. This suggests that accumulated experience in matching might characterize the method by which children learn to match. Such a position is supported by the statement that there is no evidence available to show that the teaching of discrimination is required since "the distinctive features of phonemes are not taught but they are nevertheless learned" (Gibson et al., 1962, p. 905). It appears, then, that knowledge of the relative strength of distinguishing features could affect the manner of teaching alphabet letters.

If, as Mackworth (1972, p. 716) stated, a word is maintained in the iconic store by the matching process, then presumably single letters would be similarly maintained, and the memory trace of the letter would be strengthened by the volume of matching until learning criterion of mastery is reached. When this criterion is reached may depend in part upon the child's memory storage system. As environmental information reaches the memory system, it enters primary or short-term memory. While the capacity of primary memory is limited, the process by which items pass from short- to long-term memory is considered to be under the possible control of the individual (Sticht, Beck, Hanke, James & Kleiman, 1974, p. 49). Information from the short-term memory may be retrieved accurately if it is retrieved rapidly. Speed of retrieval will depend upon the amount of preceding rehearsal and the strategies by which individuals relate new information to that already present in secondary or long-term memory. Sticht et al. (1974, p. 49) suggested that this information might be efficiently related by the linguistic recording of visually received information.

Since the capacity of primary memory is limited to approximately seven items, one item must recede as a new item enters primary memory. There may be certain conditions under which item recession is most likely to occur. Visual interference

may be one of these conditions; Barr (1972) suggested that a prereader can visually understand two to three letters at a glance, increasing by one letter per year. Since children's rates of development vary so widely, it may be that peripheral vision of nontarget letters interferes with a child's matching skill. Thus, the area of focus may control the reception of detail, and in visual discrimination it is the recognition of detail rather than the outline that ensures consistent accuracy. Geyer and Kolers (1973, p. 193) made reference to "an attentional focusing mechanism which delimits the size of the field", and it may be that training in the arousal of this mechanism could engender selective fixation, thus enabling a child to efface extraneous information.

There appears to be a need to investigate the possibility of a common surveying procedure in young children. It has long been known, for instance, that the upper half of lower-case letters is more easily identified than the lower half (Huey, 1908). More recently it has been established that stimuli in the left field of vision are more accurately perceived than those in the right field (Harcum, Hartman, & Smith, 1963), yet Kolers (1969) has shown the right side of lower-case letters to be more informational than the left side. Since the perceptual primary effect of the left field of vision appears to be strong and may be clearly differentiated from the response primary effect (Harcum, 1964), it could be that particular attention to the right side of a letter is warranted in developing skills of visual discrimination.

Hypotheses

In order to investigate the effects of different methods of training young children in the skill of discriminating letter forms, the following null hypotheses were tested.

1. Subjects who receive training in matching a stimulus with its equivalent within an array of letters show no significant difference in letter matching ability from subjects who are trained to match letters singly. (Group VS)
2. Subjects whose attention during training is drawn to a specific field of vision show no significant difference in letter matching ability from those whose attention is drawn equally to the total field of vision. (Group VS-D)
3. Subjects whose attention during training is drawn to critical cues (those particular attributes of line, curve, and directionality which differentiate letters of the alphabet) by the investigator's verbal descriptions show no significant difference in letter matching ability from subjects who receive no such direction. (Group VB-D)
4. Subjects who receive training in the oral description of critical cues show no significant difference in letter matching ability from subjects who receive no such training. (Group VB-D-R).

Design of the Study

This study was designed in four parts: screening, familiarization, training, and testing.

Subjects

The sample for this investigation included children between the ages of four years and four years six months attending day-care centres subsidized by the City of Calgary, Alberta. A total of 132 subjects were assigned in random rotating order to one of four groups. There was no control group as such, but the treatment given to Group 1 (V-S) closely approximates the usual method of teaching letter discrimination to young children.

Materials

Screening. A test of letter knowledge was given to control for prior experience. A simple test to determine the subject's understanding of the terms "same" and "different," and a test to determine the subject's initial awareness of the term "line," "curve," and "direction" were also administered.

Familiarization. Materials for this phase were constructed in exactly the same manner as were the training materials described below except that letter-like shapes were substituted for the letters used in training. Examples may be seen in Fig. 2. The purpose of this phase was to dispel any anxiety the subject may have felt, and to familiarize him with the expectations of the investigation.



Figure 2. Samples of letter-like shapes used in the study.

Training. Materials for all groups consisted of vinyl lower-case letters mounted on 8½" x 11" sheets of white paper. Letters were approximately 2" in height with a ⅜" width stroke. Twelve letters of the alphabet were used as stimulus letters in the training sessions. The letters were selected in such a way that characteristic letter features of line, curve, and direction would appear proportionately in both training and testing segments of the study. The stimulus letters were:

w t t v c z z a d d f h m p p r

The response letters were:

w a t v c y z a b d f h m g p r

For Group 1 (VS) the stimulus letter, printed in black, was placed on the left side of the page, while the response letter was one of a group of letters on the right side of the page.

For Group 2 (VS-D) the stimulus and response letters were mounted singly on opposing pages. All the letters were divided vertically in half. The left side of each letter was printed in grey, and the right side was printed in black. (Fig. 3)



Figure 3. Samples of letters used for Group 2 (VS-D).

For Group 3 (VB-D) and Group 4 (VB-D-R) materials were exactly the same. The stimulus and response letters, printed in black, appeared singly on opposing pages.

Testing Materials. All groups completed common tasks for testing purposes. For the first test (Letter Discrimination I), the following 14 lower-case stimulus letters appeared, centred singly on 8½ x 11-inch sheets:

l o i x k s y b e g j n q u.

The following lower-case response letters appeared on a further 14 sheets:

l a i x f s v b e g j m q n.

The sheets were arranged in a binder so that the appropriate stimulus/response

pairs appeared opposite each other. For the second test (Letter Discrimination II), 14 sheets were prepared, each sheet displaying 4 lower-case letters. The first letter on each sheet was the stimulus letter, while the remaining three served as response choice letters. The same stimulus letters were used as in Letter Discrimination I. For the final test (Letter Preference), a set of 18 sheets was prepared. Three different lower-case letters were spaced on each sheet. The following groups of letters were used: “o c e,” “f t h,” “d a p,” and these groups were presented in all possible orders.

Procedures

Screening, familiarization, training, and testing treatments were given individually to each subject at one sitting. Each group received a different kind of training. In all training groups, sixteen responses constituted one learning trial and each subject received two consecutive learning trials.

TRAINING GROUP	TREATMENT
Group 1 (VS)	Visual attention, undirected
Group 2 (VS-D)	Visual attention, directed
Group 3 (VB-D)	Visual attention, verbally directed
Group 4 (VB-D-R)	Visual attention, verbally directed with oral response

Figure 4. Treatments for training groups.

Training Procedures

Group 1 (VS). Subjects were asked to look carefully at the stimulus letter on the left and point to the matching letter from the group on the right side of the page. Errors were corrected as they occurred by showing the subject the correct response.

Group 2 (VS-D). Subjects were asked to look carefully at two-tone letters displayed on the opposing sides of the binder and state whether they were the same as each other. Errors were corrected as they occurred.

Group 3 (VB-D). For each subject in the third group, the binder was opened displaying two black letters on opposing pages. The subject was asked to judge if they were the same as each other. If the subject responded correctly, the investigator said, “Good. Both letters have straight lines at the left, a curve going over the top, and another straight line at the left.” The investigator’s comments varied in specific content, but always confirmed correct responses in terms of critical cues. Error responses were corrected as they occurred by drawing the subject’s attention to differences or similarities using terms of critical cues.

Group 4 (VB-D-R). For each subject in the fourth group, the binder was opened displaying two black letters on opposing pages. The subject was asked to judge if they were the same as each other. If a correct response was given, the investigator asked, “Why are they the same/not the same?” The response was recorded on a sheet designed to categorize expected responses. Twelve categories were prepared, and each category as subdivided into two columns, V and E, indicating whether the response was volunteered or elicited. If the subject’s reply did not contain reference to at least two critical differences or similarities between letters, the investigator

then asked two or three of a set of previously prepared leading questions such as, "Do they both have straight lines at the left side?", "Do they both curve over to the right?" A pencil was used as a pointer to indicate directionality. When difference or identity had been established, the subject was then asked again to say why the letters were the same/not the same. These responses were reported as elicited. If the subject's reply did contain two or more references to differences or similarities, the response was reported as volunteered. Correct responses were verbally confirmed and, if a particular point of sameness or difference had been omitted, the investigator used suitable questions from the prepared set to complete the comparison. The investigator then said, "Now tell me another reason why they are the same/not the same." Resulting responses were considered as elicited responses.

Testing Procedures

There was a common testing procedure for all subjects in all four groups immediately following the training period.

Test 1 (Letter Discrimination I). Subjects were required to compare 14 pairs of letters and state if they were the same or not the same as each other.

Test 2 (Letter Discrimination II). The subject was asked to look carefully at the letter on the left side of the sheet and point to its equivalent within the group on the right side. Fourteen responses were required.

Test 3 (Letter Preference). The subject was asked to look carefully at a group of three different letters and to select the two letters that looked most nearly the same. Three groups of letters were presented in all possible orders, thus calling for 18 responses.

Analysis of Data

A multivariate analysis of variance was used for the four types of treatment. After the MANOVA, the differences between means were investigated by tests. A transformation of raw scores was performed to determine the possible effect of letter knowledge and awareness of critical cues on the dependent variables. Data obtained from Test 3 (Preference Test) were subjected to a chi square test.

Findings and Conclusions

Tables 1 and 2 show values obtained for differences among all pairs of group means for correct responses in training trials and in Letter Discrimination I and II.

Main Effects

The treatment in which the subject was required to make an oral response in terms of line, curve, and direction was shown to be significantly superior to each of the other three treatments in developing skills in letter-matching.

Effects — Group 4 (VB-D-R). Since it was assumed that all four groups were of approximately equal learning ability, it may be posited that some of the achievement differences could be due to the attentional energy control applied by the investigator during the training of this group. Volunteered verbal responses indicated whether or not the subject was looking in the appropriate place, whether he had a purpose for looking, and whether he was conserving his attentional energy for expenditure on relevant attributes. Thus, the investigator could direct questions in the most economical and productive manner. In recording the

TABLE 1
T VALUES OBTAINED FOR DIFFERENCES AMONG ALL PAIRS OF GROUP MEANS
FOR CORRECT RESPONSES IN TRAINING TRIALS 1 and 2

Test	Treatment Group	2 (VS-D)	3 (VB-D)	4 (VB-D-R)
1	1 (VS)	1.08	3.51**	3.39**
1	2 (VS-D)		.92	2.16*
1	3 (VB-D)			1.25
2	1 (VS)	2.09*	3.12*	5.31***
2	2 (VS-D)		.31	2.79**
2	3 (VB-D)			6.34***

* $p < .05$
 ** $p < .01$
 *** $p < .001$

TABLE 2
T VALUES OBTAINED FOR DIFFERENCES AMONG ALL PAIRS OF GROUP MEANS
FOR CORRECT RESPONSES IN TESTS 1 and 2

Test	Treatment Group	2 (VS-D)	3 (VB-D)	4 (VB-D-R)
1	1 (VS)	0.0	1.49	6.99*
1	2 (VS-D)		0.665	7.21*
1	3 (VB-D)			7.07*
2	1 (VS)	0.377	0.436	5.25*
2	2 (VS-D)		0.174	5.55*
2	3 (VB-D)			6.26*

* $p < .001$

responses of these subjects, it was observed that in the second training trial there were noticeable increases in 11 of the 14 categories, and that all of these increases were volunteered as opposed to elicited responses. One of the largest increases appeared in the category recording directionality responses. All subjects in the study, except for one in the fourth group, identified "b" and "d" as being the same in the first learning trial. However, the fourth group (VB-D-R), subjects who gave an oral response, made an improvement of approximately 60 per cent on the second trial, while the second group (VS-D), subjects whose attention was directed to a specific field of vision, gained approximately 27 per cent. Improvement shown by the remaining two groups in discriminating "b" and "d" was minimal. This finding supports the view of Jackson (1972) that "confusions resulting from mirror-image type stimuli may be implicit in the stimuli and can be viewed as normal rather than abnormal," yet at the same time supports the view of Hendrickson and Muehl (1962) who suggested that research attention be directed not solely towards what children do, but towards what children can do after effective training.

Effects — Group 3 (VB-D). The subjects in this group had their attention drawn to critical cues of letters by the investigator's verbal descriptions of

similarities and differences. This group performed at a significantly higher level than the first group (VS) in both training trials, but did not maintain this superiority in letter matching. The onset of habituation may provide a possible explanation for such deterioration in performance. The condition of habituation is a normal part of the learning process, for when a particular piece of learning is complete, habituation takes over and frees the attention for the processing of new information. However, since the verbal descriptions were repetitive to some extent, they may also have become monotonous stimuli which failed to sustain the required level of alertness. If this was the case, habituation may have occurred before a sufficiently high level of learning was attained, thus accounting for the reduced number of correct responses. The sole difference between treatments for this group (VB-D) and the fourth group (VB-D-R) was that overt verbal participation was required of the fourth group. Thus it may be that the verbal response is a motivational buffer that prevents orientation from becoming habituation before the optimal time.

Effects — Group 2 (VS-D). Subjects in this group had had their attention drawn to a specific field of vision by means of an increase in the intensity of colour at the right side of the stimulus. A group of *z* scores revealed the positive effect of practice on the second training trial, followed by a dramatic decline in performance on the first letter discrimination task, where there was a 12 per cent decrease in the number of correct responses. This decrease indicates that, in training, subjects were possibly being directed towards the right field of vision, but could not function later at the same level of efficiency without the focal guidance provided by the black contrast.

Effects — Group 1 (VS). Subjects in this group had been required to match a single stimulus with its equivalent within a set of letters; these subjects achieved the lowest number of correct responses in both training and testing. The findings indicate that the subjects may have had too much opportunity for error and too little positive reinforcement for correct response. It is suggested that undirected training has insufficient transfer value for maximum learning to occur. It is also possible that the subjects' perception may have been influenced by an excess of peripheral stimuli for them to have made the appropriate selection without guidance towards relevance and purpose.

Letter Preference Test.

Subjects were asked to select from a group of three letters the two which looked most nearly alike. The response choices were "o c e," "f t h," "d a p." Chi square tests applied to the data showed preferences to be significantly beyond the chance level. In this study there were 498 selections of "c o," 736 selections of "t, f," and 535 selections of "p d," indicating that the curve in a letter, rather than its line or directionality, is a most salient feature in young children's strategies of comparison. If all subjects in the study had made the same selection with total consistency, the number of selections would have been 792. Results thus show a marked consistency of choice, the consistency being highest in the fourth group (VB-D-R). The omission rate also was highest in the fourth group, where it appeared that, for some subjects, the mind set for identity was too strongly embedded to permit a switch to similarity.

Implications of the Study

In this study it was shown that the level of letter knowledge within the groups was not a contributing factor to ability in letter matching; thus it appears that abstract expressions were used to distinguish unnamed stimuli. On the other hand, it was observed frequently throughout the study that many subjects without letter knowledge already had names for the stimuli. For instance, "m" was referred to as "the bridge," and "r" as "the gun." It is possible that some of these labels may have developed as a result of exposure to alphabet books and blocks, some of which, in an effort to make early learning attractive, obscure the critical details in decorative embellishment. Since this type of labelling is too specific to have transfer value, it is suggested that attention be paid to correct labels.

On some occasions it was noted that the subjects were comparing the centres of the stimuli rather than the outlines. Since the width of the letter stroke was $\frac{3}{8}$ inch, it appeared that this width was affecting the subject's ability to separate field from ground. Traditionally, a young child's formal introduction to the alphabet is through the use of large, thick letters. Assuming the letters are without meaning to the child, they must simply be seen as a pattern or series of markings. Hilgard (1957, p. 377) has pointed out that "patterns do not have to contain identifiable objects to be structured as figure and ground." This statement, combined with observations from the present study, would warrant questioning the practice of introducing letters only in a magnified form not usually found in reading books. With constant exposure to the magnified form, the child may be making a perception contrary to the one being reinforced by the teacher.

Observations led the investigator to believe that there exist a number of prerequisites to competency in letter matching. The child must know what a line is, and then must be able to compare lines in terms of height, slope, and relative position. He must have a spatial vocabulary in order to make meaningful reference to letter attributes and he must have a way of expressing his recognition of orientation.

The act of comparison appears to be refined in three stages: (1) recognition of gross difference, (2) recognition of fine difference, and (3) recognition of similarity. Thus, until children understand sameness, they cannot confidently express difference.

A Re-examination of Gough's Model

The model presented by Gough has demonstrated the process of the reading act in minute detail. The particular portion of the model towards which the attention of this study was directed illustrates the process by which the stimulation of the visual system forms an icon which then activates the scanner. The scanner then interrelates with pattern recognition routines before reaching the character register.

Gough introduced his article in this manner: "Suppose the eye of a moderately skilled adult reader . . ." (Gough, p. 331). Thus, he made clear his basic assumption of a degree of reading ability. However, in later discussions of the beginning reader, Gough attributed to the child all the factors necessary for "the assembly of . . . a primitive reading machine," lacking only a scanner. It is likely that all seeing people possess a scanner, and that in control of the scanner lies facility in visual discrimination. Children who have full letter knowledge have learned, within

varying degrees, to control the scanner. Children without letter knowledge must learn that control. The free scan of the human eye needs to be directed so that it functions as does an optical scanner in an electronic computer system, picking up only relevant information. The results of the present study indicate that scanner control may be facilitated through the use of oral language.

It may be noticed that Gough showed the processing sequence flowing from the scanner to the character register. Beginning readers are probably, although not necessarily, able to employ the character register as a place where letters or words appear as commonly named entities. This is not likely to happen before the child has reached the stage of letter recognition, and yet there is no reason to assume that the portion of Gough's model under discussion could not be applied to all activities of conscious looking. Therefore, it is suggested that an insertion be placed between the icon and the character register which would account for the processes involved in graphic discrimination. This proposed insertion is shown in Figure 5.

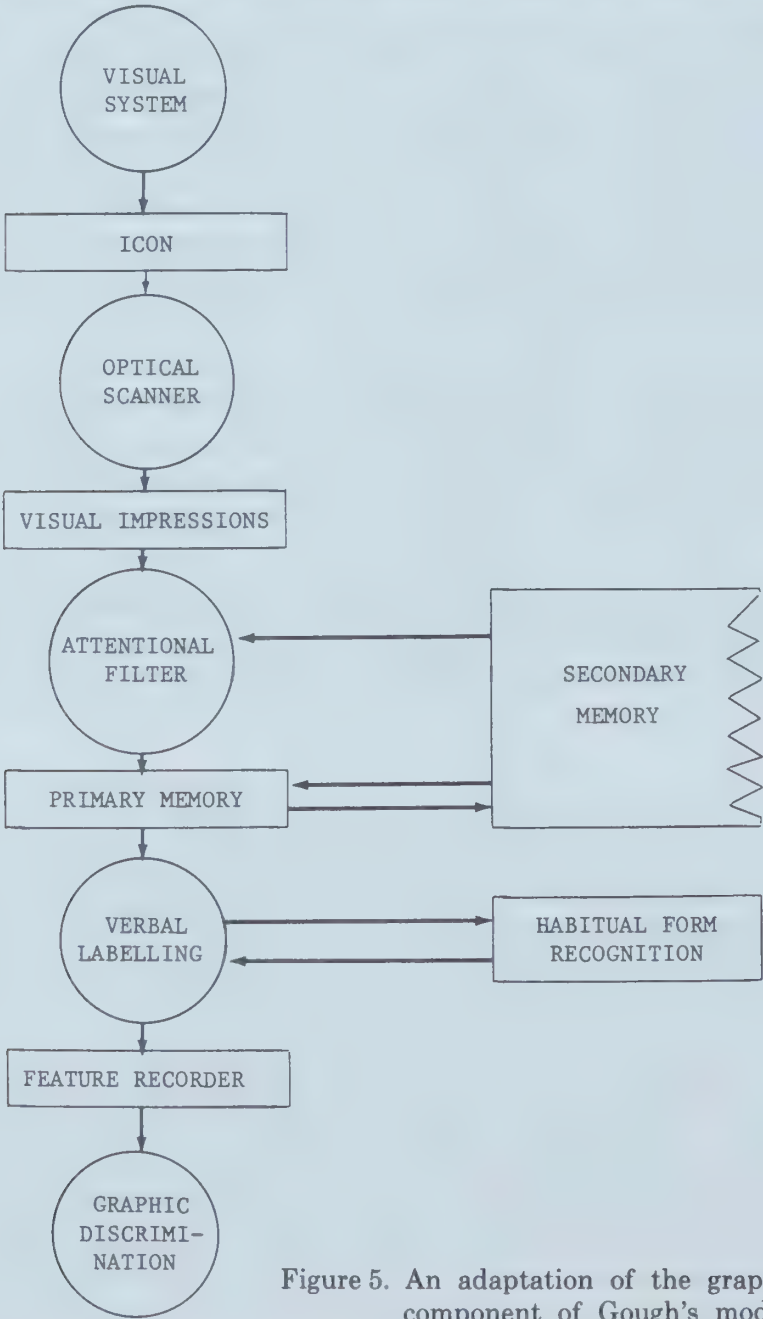


Figure 5. An adaptation of the graphic discrimination component of Gough's model of reading.

In this figure of a proposed model, the sequence moves from activation of the optical scanner to the reception of visual impressions. These impressions may be multitudinous and are of extremely brief duration. Because of their short duration the impressions must be sorted, categorized, or coded immediately in order to retain only meaningful information. Thus, a child needs to know that under given circumstances particular attributes have high relevance, and perhaps more particularly, some attributes have no relevance. In letter matching, for instance, size and colour may be irrelevant. In sorting visual impressions, a child uses his limited supply of attentional energy. Since there is a developmental aspect to the processing of attention, differences in the quality of attentional control may be expected from children of the same chronological ages. Sticht et al. (1974, p. 51) discussed developmental changes in terms of the gradually lessening degree to which the environment rules the organism. They referred to the alteration from "... the passive registration of information in displays to the active pursuing and processing of information" as the crux of the difference between seeing and looking.

Thus, much of the input of visual data is lost because it cannot be attended to before decay occurs. The visual impressions apparently pass through an attentional filter, leaving behind only those impressions which have been solidified in some manner. One of the most probable manners for solidification to occur may be through the act of labelling. By this act, features or groups of features may be clustered together under a single reference. The reference at this point need not be universally acknowledged, but must be meaningful to the individual.

The most vigorous impressions will then be taken into the primary memory. It may be recalled that Gough placed primary memory much further on in the sequence than it appears in the proposed model (Fig. 5). He defined primary memory as "a small capacity buffer storage system where four to five verbal items are maintained for a matter of seconds." (Gough, 1972, p. 339). Items within the primary memory may be ignored, or may be renewed through rehearsal. As each new item enters, it displaces an item already held within the primary memory. However, it is suggested in the proposed model that primary memory is involved in the sequence as soon as the organism becomes active beyond the initial point of visual stimulation. Gough's early placement of pattern recognition routines implies rehearsal, for routines cannot exist without rehearsal which must surely have an effect on primary memory. It is not suggested here that Gough's placement of primary memory is incorrect in the given position. The suggestion is that primary memory is active earlier in the sequence and continues to be active throughout the sequence.

Gough's (p. 340) reference to secondary memory was brief, and as a separate entity this factor did not receive prominence in his model. Smith (1971, p. 78), however, stressed that:

... nothing could be identified ... if a contribution were not made by long-term memory, because it is there that is lodged the knowledge of the world to which all incoming information must be referred.

On this basis the proposed insertion places secondary memory on a level with the attentional filter and primary memory to indicate that constant enriching interaction takes place between these factors, and this interaction may lead to habitual form recognition as a result of combining old knowledge with new. The

broken edge of the block in Fig. 5, representing secondary memory, suggests the lack of limit to the depth from which old knowledge may be drawn.

The next step in Gough's model is the character register, but nothing can be registered until it has a name. The name may be vague, perhaps as vague as "that thing" as opposed to "that other thing," for even this weak description implies a degree of discrimination. It is indicated in the proposed model that before graphic discrimination occurs, salient features of the stimuli are recorded, thus leading to discrimination. Only after discrimination would a character be registered as such. It is likely that the fineness of the discrimination will, for a young child, be related to the precision of the verbal label.

References

- Barr, R. Perceptual development in the reading process. In R. Hodges & H. Rudolf (Eds.), *Language and learning to read: What teachers should know about language*. Boston: Houghton-Mifflin, 1972, 131-139.
- Gibson, E. J., Gibson, J. J., Pick, A. D., & Osser, H. A developmental study of the discrimination of letter-like forms. *Journal of Comparative and Physiological Psychology*, 1962, 55 (6), 897-906.
- Geyer, J. J., & Kolers, P. A. Reading as information processing. In M. Voigt (Ed.), *Advances in librarianship*, 4. New York: Academic Press, 1973, 175-237.
- Gough, P. B. One second of reading. In J. F. Kavanagh & I. G. Mattingley (Eds.), *Language by ear and by eye: The relationship between speech and reading*. Cambridge, Mass.: Massachusetts Institute of Technology Press, 1972, 331-358.
- Harcum, E. R. Effects of symmetry on the perception of tachistoscopic patterns. *American Journal of Psychology*, 1964, 77 (4), 600-606.
- Harcum, E. R., Hartman, R. R., & Smith, N. F. Pre- versus post-knowledge of required reproduction sequence for tachistoscopic patterns. *Canadian Journal of Psychology*, 1963, 17 (3), 264-273.
- Hendrickson, L. N., & Muehl, S. The effects of visual discrimination pretraining with word and letter stimuli on learning to read a word list in kindergarten children. *Journal of Educational Psychology*, 1962, 52 (5), 236-241.
- Hilgard, E. R. *Introduction to psychology* (2nd ed.). New York: Harcourt, Brace, 1957.
- Huey, E. B. *The psychology and pedagogy of reading*. New York: Macmillan, 1908 (reprinted 1920).
- Jackson, M. S. Modes of adaptation to the first confrontation with English orthography in the visual modality. *Journal of Learning Disabilities*, 1972, 5(1), 29-34.
- Kolers, P. A. Clues to a letter's recognition: Implications for the design of characters. *Journal of Typographic Research*, 1969, 3 (2), 145-167.
- Mackworth, J. F. Some models of the reading process: Learners and skilled readers. *Reading Research Quarterly*, 1972, 7 (4), 701-733.
- Smith, F. H. *Understanding reading*. New York: Holt, Rinehart and Winston, 1971.
- Sticht, T. G., Beck, L. J., Hanke, R. N., James, J. H., & Kleiman, G. M. *Auding and reading: A developmental model*. Alexandria, VA: Human Resources Research Organization, 1974.
- Tansley, A. E. Readiness for reading — Can it be hastened? In M. Chazan (Ed.), *Reading readiness*. Swansea, S. Wales: University College of Swansea, 1970, 70-74.

GLEN S. AIKENHEAD

University of Saskatchewan

Using Qualitative Data in Formative Evaluation*

Formative evaluation requires sensitive techniques to assess student learning. This article explores a technique for: (a) acquiring qualitative data from standardized instruments, and (b) using these data in the formative evaluation of new curricula. Assets and limitations to the qualitative analysis technique are investigated. A new tenth grade course, Science: A Way of Knowing, was evaluated with respect to its impact on students' learning the processes, the nature, and the social aspects of science. The qualitative analysis technique proved to be a valuable tool to the developers of Science: A Way of Knowing because it yielded information that led directly to course revisions. (Dr. Aikenhead is Associate Professor in the Department of Curriculum Studies, College of Education, University of Saskatchewan.)

The general purpose of this study is to illustrate a technique for evaluating a science curriculum project. The study explores the role which qualitative data can play in formative evaluation. Assets and limitations to the qualitative analysis technique are clarified in terms of its usefulness to curriculum developers.

Qualitative data are found by asking such questions as: What ideas have students learned? What misunderstandings have they acquired or retained? The essential advantage to using qualitative achievement data is in their power to reduce ambiguity in the evaluation process.

Test scores can be very ambiguous. The following claim is typical: "Students of course E gained 3.56 points on test T while students of course C gained only 1.71 points." These gains may be statistically significant, but they give the curriculum developer and teacher no indication as to what students actually learned. Test scores by themselves have very limited meaning (Aikenhead, 1973; Lucas, 1975).

Qualitative data are derived by analyzing test items and noting the content of those items. The items may be selected simply because of a high number or a low number of correct responses they received from students. (Examples are cited below.) For a pretest posttest design, one can identify items for which students

*Based on a paper presented to the 51st annual meeting of the National Association for Research in Science Teaching, Toronto, Canada, April 2, 1978.

made a significant gain or loss during a period of instruction. Consequently, student gains in understanding and in misunderstandings are identified. This type of item analysis is accomplished with the McNemar (1969) chi squares statistic. Again, examples are cited below.

Only a few studies have taken advantage of the articulation of student learning generated by qualitative data. Cooley and Klopfer (1963) first demonstrated its effectiveness when they evaluated Klopfer's *History of Science Cases*. The McNemar chi square indicated test items which showed significant student gains. The content of those items comprised the qualitative data needed for the summative evaluation of the case studies. Klopfer and McCann's evaluation of a unified science course lent further credence to the usefulness of qualitative data. They identified weak areas in student understanding and made specific recommendations to teachers concerning what needed improving in the course (for example, "the aims and activities of science"). MacKay (1971) studied the common misunderstandings among seventh to tenth grade students in Australia. He was able to indicate which false notions seemed to be unchanged over those years. Similarly, Tamir (1972) documented some major misconceptions which Israeli high school students held about the nature and processes of science. It was hoped that these qualitative data would influence the science instruction in Israel. *Project Physics* was evaluated in terms of knowledge acquired about science and scientists (Aikenhead, 1974b). The McNemar chi square item analysis pointed to a specific knowledge for which students showed a significant increase or decrease in understanding. Thus the *Project Physics* and control groups were compared in a manner amenable to teacher interpretation. For example, *Project Physics* students generally gained in their understanding of the tactics of science, its values, its institutional functions, and the interaction of science with society, while the control group made almost no improvement in these areas at all. In the same arena of knowledge, Aikenhead (1975a) evaluated the Saskatchewan Chemistry program by: (1) listing specific misconceptions that were generally held by chemistry students, and (2) assessing the program's impact on students' knowledge about science and scientists. This assessment relied on qualitative data generated by the McNemar statistic. The findings of that study led to a major revision in the province's chemistry syllabus (Aikenhead, 1977).

In the summative evaluation projects described above, qualitative data played a supporting role. However, qualitative data could likely serve a more crucial function in the evaluation process when used in formative evaluation (Aikenhead, 1974b). This function has not been critically explored. The present study investigates the usefulness of qualitative data in formative evaluation.

Procedure

Treatment

A new kind of science course, *Science: A Way of Knowing* (Aikenhead & Fleming, 1973), is presently being developed in Saskatchewan, Canada, for tenth grade students of average ability. The full year course is an activity-centered, interdisciplinary approach to learning *about* science and scientists: the aims of science, its epistemology, its tactics and strategies, its values, its interaction with society, and its human characteristics. *Science: A Way of Knowing* allows for a high degree of teacher flexibility in the content, sequence, and instructional methods (Aikenhead, 1975b).

Field Trials

In 1974, four Saskatchewan teachers (three urban, one rural, total number of students = 141) volunteered to teach *Science: A Way of Knowing* draft #2. Feedback was gained through parent and student questionnaires, teacher and student interviews, and achievement tests. As a result, draft #2 was polished by making a number of small changes such as using appropriate cartoons, rewriting some thought provoking questions, and adding more optional activities.

In 1975, four teachers field tested draft #3 (two urban, one rural, and one — the author — in a Swiss International school). Only one teacher (urban) participated in both field trials. The total number of students was 180.

The study reported here is part of the formative evaluation of *Science: A Way of Knowing* draft #3. The developers wished to know:

- a. What ideas have students learned *about* science and scientists?
- b. What misunderstandings have they still retained?
- c. How do the answers to the above questions vary when the course is taught by different teachers?

These questions lend themselves to a one-group pretest posttest design using the McNemar chi square item analysis. At this stage in the development of *Science: A Way of Knowing*, there is no intention of comparing treatment groups with control groups. The teacher effect on student achievement is taken into consideration through the comparison of draft #3 results with draft #2 results.

Instruments

To assess knowledge *about* science and scientists, students wrote the *Science Process Inventory*, SPI (Welch, 1969), and the *Test on the Social Aspects of Science*, TSAS (Korth, 1968). A detailed description of these instruments, including their validation and their use in research, may be found elsewhere (Aikenhead, 1973). The reader is encouraged to note some of the philosophical assumptions inherent in these instruments, as described by Lucas (1975).

The SPI deals with a student's awareness of the activities, assumptions, products, and ethics of science. It has 135 items. The form of the SPI usually used by researchers, form D, has a forced choice response format. Students must agree or disagree with each item. In an earlier study evaluating *Project-Physics*, Aikenhead (1972, p. 221) found that the forced choice format led to ambiguities for the researcher. One did not know if a student was disagreeing with a statement because: (a) she really did disagree, or (b) she did not understand what the item meant. Therefore the agree/disagree format of form D was replaced by a Likert-type scale: 1. strongly agree, 2. agree, 3. do not understand the meaning, 4. disagree, and 5. strongly disagree. The response "do not understand the meaning" allows students to indicate that a choice is impossible due to vocabulary or syntax. A student is forced to take a stand so long as she thinks she understands the meaning of the item. In order to distinguish between the two response formats, the Likert-type response instrument will be called "SPI form DL." Because students are not forced to guess with the Likert-type scale, the average scores should be lower for form DL.

The TSAS deals with: the interactions among science, technology and society; the social nature of the scientific enterprise; and the social and political responsibilities of scientists. There are 52 items. The original five-point Likert

scale was modified slightly. The mid response “uncertain or undecided” was changed to “do not understand the meaning.” Again, this was done to force students to take a stand if they felt they understood the item.

Both tests were marked in the same manner. Each item has a “correct” response based on the content validity of the instrument. The response “agree” or “strongly agree” or the response “disagree” or “strongly disagree” was considered correct if it concurred with the test’s key. The “do not understand the meaning” response was always marked incorrect, however it was not ignored. Each item was later analyzed with respect to the percentage of students who agreed with, did not understand, or disagreed with, the item.

Data Analysis

The usual test score information was generated in two ways: (1) for each of eight teacher groups, and (2) for each of two combined groups, draft #2 and draft #3. A *t*-test for matched pairs determined the significance of gains between the pretest and posttest scores, for both the SPI form DL and the TSAS.

Using the draft #3 combined group data, each test item was analyzed by the McNemar (1969) chi square statistic.

$$\chi^2 = \frac{(A - D)^2}{(A + D)}$$

with 1 degree of freedom.
where A and D are cell frequencies
in the following contingency table.

		Posttest	
		right	wrong
Pretest	wrong	A	B
	right	C	D

This statistic identified test items which experienced a statistically significant increase or decrease in student response between the pretest and posttest.

Each test item was also analyzed to determine the percentage of students making one of the three responses: agree, do not understand the meaning, and disagree. (The percentage of students not responding at all was also noted.)

Results and Findings

This report deals with the formative evaluation of *Science: A Way of Knowing*, draft #3. Occasionally, information from draft #2 is included for reasons of comparison only. Student responses to the SPI form DL and TSAS were analyzed both quantitatively and qualitatively.

Quantitative Analysis

Table 1 contains standard data for the SPI form DL and TSAS scores. The gains between the pretests and posttests are statistically significant. Students’ achievement was slightly greater for the SPI form DL than for the TSAS. (The actual gains are particularly large in comparison with others reported in the literature.) In spite of a change of teachers and in spite of an “improvement” in the instructional materials, the gross total test scores seem stable over drafts #2 and #3.

TABLE 1
STUDENT TEST RESULTS FOR TWO DIFFERENT FIELD TRAILS,
TEACHER GROUPS COMBINED

	Draft #2 — 1974-1975						Draft #3 — 1975-1976					
	SPI Form DL (N=137)			TSAS (N=109)			SPI Form DL (N=171)			TSAS (N=149)		
	Pre- test	Post- test	Gain	Pre- test	Post- test	Gain	Pre- test	Post- test	Gain	Pre- test	Post- test	Gain
Mean score ^a	85.0	96.0	11.0	27.4	31.1	3.7	85.4	96.5	11.1	28.0	31.7	3.7
Standard deviation	13.1	12.4		5.4	7.1		11.3	12.2		5.4	6.7	
Reliability ^b	.85	.85		.65	.78		.80	.84		.67	.76	
t test (matched pairs)	10.9	p < .001		5.4	p < .001		15.2	p < .001		8.92	p < .001	

^a Total possible scores are: SPI, 135; TSAS, 52.
Typical mean scores for Saskatchewan 11th grade Chemistry students are: SPI Form DL, 92; TSAS, 28.
^b K-R formula - 20 coefficient.

Information based on teacher groups is shown in Table 2. There is variation among the teacher groups, especially for the TSAS scores. However, other studies reveal even greater variation among teacher groups (Aikenhead 1972, 1975a). These results support the cliché that the individual teacher has more effect on student achievement than a particular curriculum. (The idea behind *Science: A Way of Knowing* is to give support to teachers having interests which correspond to the objectives of the course.) The predominant “no significant difference” for TSAS total scores suggests that students are not learning much about the social aspects of science. However, this is where quantitative data are most limited in giving useful information to a curriculum developer.

TABLE 2
SUMMARY OF STUDENTS' TOTAL SCORES GROUPED BY TEACHER

Teacher	SPI Form DL						TSAS					
	N	Pretest	Posttest	Gain	t-test ^a	Proba- bility	N	Pretest	Posttest	Gain	t-test	Proba- bility
Draft #2												
A	76	86.1	98.4	12.3	10.4	p<.001	43	26.3	43.1	7.8	8.1	p<.001
B	21	77.2	91.0	13.8	5.6	p<.001	22	26.0	27.0	1.0	0.8	not sig.
C	16	88.6	98.5	9.9	2.7	p<.02	15	31.4	34.4	3.0	1.8	not sig.
D	24	85.7	91.0	5.3	1.7	not sig.	29	28.0	28.0	0.0	0.0	not sig.
Draft #3												
A	50	86.7	101.8	15.1	12.8	p<.001	45	29.1	36.5	7.4	11.0	p<.001
E	16	83.9	98.7	14.8	6.3	p<.001	— ^b	—	—	—	—	—
F	69	84.3	94.2	9.9	8.7	p<.001	68	27.8	30.7	2.9	5.2	p<.001
G	36	86.2	92.5	6.3	4.3	p<.001	36	26.9	27.7	0.8	1.2	not sig.

^a t-test for matched pairs.
^b No data were available.

Qualitative Analysis

The gain scores, statistically significant or not, do not shed any light on what ideas students tend to learn in their study of *Science: A Way of Knowing*. The curriculum developers wished to know what course content was generally successful and what content required improvement. The McNemar chi square item analysis identified specific ideas for which students showed significant increases or decreases in understanding. Table 3 summarizes the achievement for both draft #2 and #3 field trials. A significant increase or decrease in a correct response between the pretest and posttest is indicated by a "+" or "-" respectively. Although Table 3 shows no differentiation among levels of probability, it may be worth noting that for the draft #3 group: (a) $p < .001$ for 38 items; (b) $p < .01$ for 22 items; and (c) $p < .05$ for 29 items. The test items in Table 3 are categorized by their content. The arbitrary classification scheme is purely a heuristic one.

Even though there is some variation in student achievement between drafts #2 and #3, the achievement of the two groups generally appears to be quite consistent. The few inconsistent areas are: tentativeness of scientific knowledge in general, the value of opinion versus fact, inquiry tactics used by scientists, assumptions, and acceptance of scientific results. (The latter two ideas did experience some improvement in draft #3.) It is interesting to compare these results with those of another study involving a unified science course (Klopfer & McCann, 1969). The consistency in all student achievement between successive years in the Klopfer and McCann study was equivalent to the consistency between successive years in those areas cited above as being "inconsistent." Thus, draft #2 and #3 groups appear to be very similar in the ideas they learned. Considering that the total gain scores varied widely among teacher groups (Table 2), it would seem that student achievement takes place within a fairly consistent cluster of items. In other words, Table 3 lists those items which generally meet with success regardless of the total *quantitative* gain of the group; that is, regardless of the teacher. Therefore, while different teachers appear to influence the quantitative gain in achievement, they influence the qualitative content of that student achievement much less for *Science: A Way of Knowing*. This is useful information for the developers of the course.

Given this fairly consistent student achievement among different teachers, the germane question becomes: *What ideas* do students learn about science and scientists? Superficially this may be answered by noticing the descriptive titles for the items which show gains during the year (Table 3). However, the question is better answered in depth by carefully examining the content of the items clustered together in Table 3.

Because the major purpose of this study is to explore the value of such information, and because space does not allow for a complete analysis here, four clusters of items have been selected for detailed analysis (aims of science, role of opinion and fact in science, "the scientific method" myth, and the role of classification in science). The four clusters were selected because they exemplify the various assets and different problems encountered in the qualitative analysis technique.

The results compiled in Table 3 indicate that students learned about the aims of science. In the past, several surveys had indicated that students confuse science with technology, especially the principal aim of each endeavour (Korth, 1969; Hurd, 1969; Mitias, 1970). In the course being evaluated, a unit entitled

TABLE 3
SUMMARY OF ITEMS EXPERIENCING A SIGNIFICANT CHANGE
BETWEEN PRETEST AND POSTTEST^a

Topic	Number of Items		Draft #2	Draft #3
	SPI	TSAS		
<u>Aims of Science</u>	3	4	+ ^b	+
		1	-	-
<u>Epistemology of Science</u>				
Assumptions	4		+	+
	1		+	
	5			+
	2		-	
General Tentativeness versus Truth	1		+	+
	2		+	
		1	+	+
		1		+
Observations				
	2		+	+
Measurements	2		+	+
	2			+
Hypotheses	4		+	+
	1		+	
Laws	3		+	+
	2			+
	1		+	
Theories	2		+	+
	1			+
	1		-	
Models	7		+	+
	1			+
<u>Tactics of Science</u>				
"The Scientific Method" Myth	1	1	+	+
	1			+
Classification	3		+	+
	1			+
Inquiry	6		+	+
	2		+	
	2			+
	1		-	
Induction/Deduction	2		+	+
	1			+
<u>Values Within Science</u>				
Simplicity	3		+	+
Acceptance of Results	1		+	+
	4			+
	2		+	
Quantification	1		+	+
Practicality	1	2	+	+
		1	+	
Open Communication	1		+	+
		1		+
<u>Interaction of Science With Society</u>		2	+	+
		2		+
		1	+	
		1	-	
<u>Human Characteristics</u>		3		+
		3	+	

^a A detailed identification of individual items is available from the author.

^b A statistically significant McNemar chi square value (0.05 level of probability) is signified by the symbols "+" for an increase in correct response and "-" for a decrease in correct response.

“Technology: A Way of Knowing” distinguishes between a technologist (providing products and techniques for public use) and a scientist (satisfying his or her curiosity about the natural world). While this dichotomy is as simplistic as “pure” and “applied” science, it was nevertheless useful in bringing some order out of

TABLE 4
ITEMS CONCERNING THE AIMS OF SCIENCE

Items		Correct Response		% Response			χ^2 ^b	p <
				agree	? ^a	disagree		
T9	The primary objective of the working scientist is to improve human welfare.	disagree	pretest posttest	65.6 39.6	4.7 0.7	29.7 59.7	33.2	.001
T22	The principal aim of science is to provide the people of the world with the means for living better lives.	disagree	pretest posttest	75.8 40.0	1.3 2.1	22.8 57.9	32.9	.001
T24	Science is primarily a method for inventing new devices.	disagree	pretest posttest	33.8 23.6	3.4 0.7	62.8 75.7	7.1	.01
S129	The primary objective of science is to develop new and improved living conveniences.	disagree	pretest posttest	56.2 31.5	3.0 0.0	40.8 68.5	26.5	.001
T34	The greatest accomplishments of science consist of the many useful products it has produced.	disagree	pretest posttest	73.6 47.3	4.1 4.1	22.3 48.6	25.9	.001
S86	One of the aims of science is to work towards more complex knowledge.	disagree	pretest posttest	75.4 62.0	5.3 0.6	19.3 37.4	16.9	.001
S94	Prediction is an important goal of scientific investigation.	agree	pretest posttest	67.5 77.8	3.6 1.8	29.0 20.5	6.6	.05
T43	The aim of science is to increase man's knowledge of the physical and biological world.	agree	pretest posttest	88.5 78.4	2.7 1.4	8.8 20.3	5.5	.05
S51	The main object of basic scientific research is the discovery of understanding rather than its practical application.	agree	pretest posttest	64.1 71.3	20.6 12.3	15.3 16.4	2.7	NS ^c

^a ? means "do not understand the meaning"
^b McNemar chi square
^c NS means "not significant"

chaos in the minds of adolescents. Later in the course the two extremes are integrated in a less naive fashion in the context of the everyday world. (The philosophical stance that science searches for truth was rejected by the authors of *Science: A Way of Knowing*. The course basically embraces the Thomas Kuhn approach to understanding science.) The specific items listed in Table 3 under “aims of science” may be found in Table 4. The first four actually refer to a principal aim of technology. Many students seem to learn that science is *not* technology but do they learn what science is? TSAS item 43 and SPI item 51 are the only ones in both instruments which state a major aim of science. The decrease in learning reflected by item TSAS 43 and the lack of a more positive result on item SPI 51 may be explained by the result that students rigidly focussed on the curiosity aim, and therefore they do *not* translate that idea into human behaviours such as those described in item TSAS 43 and SPI 51. Therefore, draft #4 of *Science: A Way of Knowing* will have to require that students explore how the curiosity aim manifests itself in a number of ways in university and industrial laboratories. Not only does one section in the technology unit need to be rewritten, but several parts throughout the course will need modifying to reinforce the idea.

Table 3 contains a blatant example of misunderstandings apparently being learned (subsection “opinion versus fact”). A closer analysis yields useful information for the curriculum developer and evaluator (see Table 5). In order to understand how students interpret some of these items, a group not involved in the study was asked to write a justification for their responses. By reading student written responses, one finds that students tend to view SPI item 41 as simplistic and therefore reject it on that account. They fail to see its underlying sophisticated

TABLE 5
ITEMS CONCERNING THE ROLE OF OPINION AND FACT IN SCIENCE

Items		Correct Response	% Response			χ^2	$p <$	
			agree	?	disagree			
S41	Science must start with facts and end with facts no matter what theoretical structures it builds in between.	agree	pretest posttest	55.0 39.4	6.4 8.2	38.6 52.4	11.6	.001
S44	The assumptions in science are based on opinion, not fact.	disagree	pretest posttest	50.3 65.7	2.3 1.2	47.4 33.1	9.1	.01
S60	Scientific conclusions should be based on facts, not opinion.	agree	pretest posttest	84.8 74.1	0.6 0.0	14.6 25.9	6.6	.05
T45	The scientific investigation of human behaviour is of little value since it must involve the personal opinion of the investigator.	disagree	pretest posttest	39.0 28.4	9.6 10.1	51.4 61.5	3.9	.05

meaning. On the other hand, the “correct” affirmative student responses encompass a wide range of justifications, ranging from naive to sophisticated. Quite differently, SPI items #44 and #60 (Table 5) represent an interesting problem for the developers of *Science: A Way of Knowing*. In the unit “An Artistic Way of Knowing: Perception,” students become involved with the role of basic assumptions in perception. Students study the influence that culture in general and education in particular have on these assumptions. They learn that different educational backgrounds likely lead to different perceptions or world views. Apparently this relativism encourages students to associate opinion with a world view; “world views are culturally relative thus they are in the realm of opinion, not fact,” they reason. Students consequently believe that a scientist’s world view (perceptions, assumptions, paradigm) is based on opinion. However, it seems clear that the SPI items associate “fact” with “world view” and consequently for these items “opinion” has the connotation of conjecture. This connotation is not shared by students who have adopted the relativism position inherent in the course. One can discern this relativism in item TSAS #44 (Table 5). Understandably, there is a significant increase in correct response for this item. The interesting problem which emerges from this analysis is that “negative” learning (SPI items #44 and #60) is actually “positive” learning (TSAS item #45). The definition of what is negative and what is positive relies on one’s philosophical stance (Lucas, 1975). Curriculum developers must be very aware of the curriculum’s philosophical stance so that they can properly interpret student achievement in terms of that stance (Campbell, 1971). In the case of *Science: A Way of Knowing*, the “negative” test results for SPI items #44 and #60 do not seem to warrant any specific changes in the curriculum materials. The apparent misunderstanding over opinion versus fact turns out to be more an artifact of the test itself.

At the end of the course, students believed less in “the scientific method” as a step by step procedure which: advances knowledge, produces correct answers, and leads to successful experiments (Table 6). However, a majority of students on the posttest (55%) still expressed the belief that there was such a method. The developers may not be satisfied with the dramatic positive jump from 25% to 43% in correct response. If they are not, then many places in the student’s manual will need to be rewritten in a way that challenges the application of “the scientific method.” Already in the course there are questions and discussions dealing with the topic. More would seem justified.

TABLE 6
ITEMS CONCERNING “THE SCIENTIFIC METHOD” MYTH

Items		Correct Response		% Response			χ^2	$p <$
				agree	?	disagree		
T3	Scientists have advanced knowledge by consistently following, step by step, a definite procedure called the scientific method.	disagree	pretest posttest	69.6 55.4	5.4 2.0	25.0 42.6	10.9	.001
S84	The scientist knows that his experiment will be successful if he follows the steps of the scientific method.	disagree	pretest posttest	32.4 22.9	2.4 0.6	65.3 76.5	7.1	.01
S124	Applying the scientific method to a problem will always produce the correct answer.	disagree	pretest posttest	20.4 11.2	2.4 0.6	77.2 88.2	9.8	.01

Students made major gains in their appreciation of classification schemes in science (Table 7). It is interesting to note that SPI item 61 is for the most part not understood. From reading students’ written justifications, it is apparent that the word “inherent” is the major reason for this. The developers of *Science: A Way of Knowing* apparently need not make any alternations regarding the nature of classification schemes in science.

TABLE 7
ITEMS CONCERNING THE ROLE OF CLASSIFICATION IN SCIENCE

Items		Correct Response		% Response			χ^2	$p <$
				agree	?	disagree		
S52	A classification scheme, such as the periodic table of the elements, is based on common factors and differences noted in observations.	agree	pretest posttest	47.4 73.1	43.3 15.8	9.4 11.1	22.5	.001
S61	Classification schemes are inherent in the materials classified, rather than imposed on nature by the scientist.	disagree	pretest posttest	27.2 30.4	62.7 49.7	10.1 19.9	7.8	.01
S78	Classification schemes are a useful means of organizing observations.	agree	pretest posttest	63.5 87.1	27.1 7.1	9.4 5.9	26.7	.001
S83	Grouping observations is an important part of scientific work.	agree	pretest posttest	81.2 90.0	7.6 1.8	11.2 8.2	6.8	.01

Discussion

The formative evaluation process can be greatly enhanced by using qualitative data. As seen in the few sample analyses above, the developers of *Science: A Way of Knowing* received information which had a direct application to revising their materials and activities. Qualitative data are substantial valid sources of feedback in formative evaluation.¹

The usefulness of qualitative data can be clarified by noting what information an evaluator receives when his work goes beyond the usual quantitative analysis.

The quantitative data generated in this study reflected a significant increase in students’ understanding about the processes of science. However, the *qualitative* data clearly indicated what these understandings were: assumptions underlying science; the nature of observations, measurements, hypotheses, laws, theories, models, and classification schemes; the mythology of “the scientific method”; the notion of induction and deduction; and values related to simplicity, acceptance of results, quantification, and practicality of scientific results.

Unlike the quantitative analysis, the qualitative analysis was able to reveal some apparent misunderstandings inherent in the course materials. In most cases, however, the philosophical positions between the test items and the course objectives were different. These apparent misunderstandings can be dealt with in the appropriate way by the course developers.

The quantitative analysis clearly suggested that *Science: A Way of Knowing* does not generally address itself to ideas about the social aspects of science. However, the *qualitative* data delineated ideas which were generally learned: the distinction between science and technology, the relationships between the two, the importance of open communication to the advancement of science, the tentativeness of scientific knowledge, some human characteristics of scientists, and the role of science in helping to solve social problems. Knowing this, the developers can make changes that simply emphasize these ideas, thereby affecting the quantitative achievement of students. At the same time, completely new material that may affect the qualitative achievement would not seem necessary for these same ideas. On the other hand, if there are item clusters which show no significant improvement, and if the item clusters relate to a desired goal of the course, then the developers may need to construct new activities and materials.

The quantitative data indicated that student achievement varied with the teacher. However the qualitative data showed a high degree of consistency among the teacher groups in terms of what content was learned. This finding is relevant to inservice training for *Science: A Way of Knowing*. Once teachers have a consistent philosophical posture with respect to the major content of the course (as these teachers seem to have had), some will need to develop teaching techniques in order to emphasize this content.

The quantitative data gave no indication of the knowledge deficiencies students had at the end of the course. On the other hand, the qualitative data explicitly showed that there was still room for improvement in such areas as: the analysis of an experiment in terms of its control and its variables, the tentativeness of scientific laws, a lingering belief in "the scientific method," and the unwarranted faith in scientific beliefs and attitudes outside the realm of scientific inquiry. These items experienced less than 50% correct student response.

By noting the proportion of students who thought they did not understand the meaning of an item, the evaluators were able to pinpoint specific reasons why certain items were not correctly answered. The wording of some SPI items turned out to be particularly difficult (not understood by more than 25% of the students); for example: #43 ("experimental control"), #57 ("control of variable"), #61 ("inherent"), #120 ("tentative"), and #132 ("degree of estimate"). This feedback illustrates a further reward from using qualitative data. An evaluator ought to be able to distinguish between students' misunderstanding an idea and their not understanding an item's wording.

Certainly the content of the SPI and TSAS does not encompass all the major objectives of the course. The ones excluded must be evaluated in some other way (Champagne & Klopfer, 1974). On the other hand, there is likely to be content in the SPI and TSAS which falls outside the objectives of *Science: A Way of Knowing*.² Nevertheless, it seems advisable for a curriculum developer to assess the impact of a course beyond the parameters of its specific objectives (Aikenhead, 1974a). Klopfer (1969) clarified this issue when he described the *effectiveness* of a course in terms of how closely it met its objectives, while its *effect* embraced a

wider range of outcome. For this reason, it was worth analyzing all the items on the SPI and TSAS, whether or not the developers perceived them as relevant to the course. If the course is having positive or negative effects unanticipated by the developers, they should be aware of these.

In a number of instances, the SPI and TSAS seemed to assess a student's *sophistication* in knowledge about science and scientists. This phenomenon seems important to recognize when using qualitative data. For the purpose of the present discussion, one can simply assume that from *ignorance or confusion* a student develops a *basic* understanding and then from there a *sophisticated* understanding. Before she can become sophisticated in an area of knowledge, she must first have a basic, perhaps simplified, understanding. Many of the Saskatchewan tenth grade students harboured misunderstandings about science and scientists which had to be improved to simple understanding before sophistication could be achieved. A number of items from the SPI and TSAS would seem to reflect an improvement from simple understanding to sophisticated understanding, while other items reflect an improvement from misunderstanding to simple understanding. Therefore over a one-year period, there may be a little improvement on a number of items of sophistication because students are involved with correcting misunderstandings. This might explain the results concerning the "aims of science." Students tended to correct a misunderstanding (the confusion of science with technology) but they were not able to think of the principal aims of science in the sophisticated way anticipated by certain test items (Table 4). Conversely, when students have a *basic* understanding of some idea, one would not expect to see much improvement on relevant items of basic understanding. However one would hope to notice significant improvement on items assessing sophistication in that same idea.

While qualitative data provide a curriculum developer with feedback having much less ambiguity than quantitative data, qualitative data naturally have limitations. Because students bring their subjective interpretation to each item, their response on a Likert five-point scale has inherent subjectivity. In a number of instances, the developers of *Science: A Way of Knowing* could only interpret student responses by having students write a justification for their responses. These written responses revealed a range of misunderstandings as well as a range of understandings — from simple to sophisticated. This written information gave some insight into the subjectivity of "objective" tests. It also lent considerable insight into the impact of the course. Because students related to the course content in idiosyncratic ways, their written responses often exposed these idiosyncracies and allowed the reader some opportunity to assess the kind of impact the instruction had had. *Future work in formative evaluation might devise some fairly reliable techniques for achieving an in-depth analysis of student learning.*

The qualitative analysis procedure proved to be a valuable tool in the repertoire of techniques used in formative evaluation. The role of qualitative data appears to be more extensive in formative evaluation than it is in summative evaluation, a prediction which prompted this study (Aikenhead, 1974b).

Notes

1. Other sources include: student achievement on tests constructed specifically for the course content; student achievement in activities directly related to those of the course; student achievement in analyzing science and technology content of the mass media; parental reaction to their child's work; the effect on student achievement in subsequent classes, not only

science classes, but other subject areas as well; the effect on the school from having an interdisciplinary offering. For a complete analysis of the formative evaluation process, see Champagne and Klopfer, 1974.

2. The TSAS, for instance, contains a number of items related to a scientist's ethical responsibility to society. *Science: A Way of Knowing* does not focus on this issue; it is left to a subsequent course where students are more mature to handle this content.

References

- Aikenhead, G. S. The measurement of knowledge about science and scientists: An investigation into the development of instruments for formative evaluation. Unpublished doctoral thesis, Harvard University, Cambridge, Mass., 1972.
- Aikenhead, G. S. The measurement of high school students' knowledge about science and scientists. *Science Education*, 1973, 57, 539-549.
- Aikenhead, G. S. Course evaluation. I: A new methodology for test construction. *Journal of Research in Science Teaching*, 1974, 11, 17-22. (a)
- Aikenhead, G. S. Course evaluation. II: Interpretation of student performance on evaluative tests. *Journal of Research in Science Teaching*, 1974, 11, 23-30. (b)
- Aikenhead, G. S. Learning about the scientific enterprise in Saskatchewan high school chemistry classes: An evaluation. *Saskatchewan Journal of Educational Research and Development*, 1975, 5, 4-29. (a)
- Aikenhead, G. S. A new kind of science experience in the high school. A paper presented to the 23rd National Convention of NSTA, Los Angeles, March, 1975. (b)
- Aikenhead, G. S. New developments in Saskatchewan' chemistry curriculum. A paper presented to the 17th Annual Conference of the Science Council A.T.A., Banff, Alberta, Canada, October 9, 1977.
- Aikenhead, G. S., & Fleming, R. W. *Science: A way of knowing*. Saskatoon, Sask.: Department of Curriculum Studies, University of Saskatchewan, 1973.
- Campbell, D. C. Epistemological posture as an intellectual variable. Ontario Institute for Studies in Education: Department of Curriculum, 1971 (mimeographed).
- Champagne, A. B., & Klopfer, L. E. Formative evaluation in science curriculum development. *Journal of Research in Science Teaching*, 1974, 11, 185-203.
- Cooley, W. W., & Klopfer, L. E. The evaluation of specific educational innovations. *Journal of Research in Science Teaching*, 1963, 1, 73-80.
- Hurd, P. *New directions in teaching secondary school science*. Chicago: Rand McNally & Co., 1969.
- Klopfer, L. E. Effectiveness and effects of ESSP astronomy materials. *Journal of Research in Science Teaching*, 1969, 6, 64-75.
- Klopfer, L. E., & McCann, D. C. Evaluation in unified science. *Science Education*, 1969, 53, 155-164.
- Korth, W. W. The use of the history of science to promote student understanding of the social aspects of science. Unpublished doctoral thesis, Stanford University, 1968.
- Korth, W. W. Test every senior project: Understanding the social aspects of science. A paper presented at the 42nd meeting of the National Association for Research in Science Teaching, Pasadena, California, February 8, 1969.
- Lucas, A. M. Hidden assumptions in measures of "knowledge about science and scientists." *Science Education*, 1975, 59, 481-485.
- MacKay, L. D. Development of understanding about the nature of science. *Journal of Research in Science Teaching*, 1971, 8, 57-66.
- McNemar, Q. *Psychological statistics* (4th ed.). New York: John Wiley, 1969.
- Mitias, R. G. E. Concepts of science and scientists among college students. *Journal of Research in Science Teaching*, 1970, 7, 135-140.
- Tamir, P. Understanding the processes of science by students exposed to different science curricula in Israel. *Journal of Research in Science Teaching*, 1972, 9, 239-245.
- Welch, W. W. *Welch Science Process Inventory, Form D*. Minneapolis, Minn.: Dr. Wayne Welch, 1969.

BOOK REVIEWS

RESEARCH INTO TEACHING METHODS IN HIGHER EDUCATION, 4th Edition. By Ruth M. Beard, Donald A. Bligh and Alan G. Harding. Guildford, Surrey: Society for Research into Higher Education Ltd., 1978, 150 pp. (ISBN 0-900868-58-9)

The fourth edition of this monograph is an updated and somewhat enlarged version of one which appeared first in 1967. The Society for Research into Higher Education serves universities, polytechnics and colleges of education in Great Britain; this monograph is one part of that program of service focussed on British problems. One notable quotation from the Introduction provides a partial background:

Until comparatively recently all changes in university teaching were due to outstanding innovators in the universities, or followed on recommendations of committees and professional bodies. Few of these have been directly influenced by findings in the psychology of learning or experiments into the effectiveness of teaching methods; they were based almost exclusively on teachers' views as to how the subject should develop, their experience of learning and teaching and knowledge of methods used elsewhere. (p. 2)

For people interested in the improvement of instruction in colleges, technical institutes and universities, the Beard et al. work is a useful review. However, because of its deliberate focus upon British sources to the exclusion of North American and other sources, it suffers from a special kind of flatness of perspective.

A list of chapter headings may provide potential readers with some clues about the relevance of the monograph to their own research and teaching interests:

Chapter 1: Aims and Objectives; Chapter 2: Economy and Efficiency; Chapter 3: Recall and Retention of Information (including fairly lengthy sections on: programmed learning, the lecture, and audio-visual aids); Chapter 4: Skills and Abilities (including: laboratory skills, study skills, oral skills and group discussion, and higher mental skills); Chapter 5: Teaching for Change of Attitudes; Chapter 6: Evaluation of Students, Teachers, and Teaching Methods.

The bibliography contains over five hundred entries and is, in itself, a valuable set of information. Perhaps the monograph would benefit from the newly-developing meta-analytic approaches to research summaries which Glass (Glass, Gene V., "Primary, Secondary, and Meta-Analysis of Research" in *Educational Researcher*, November 1976) and some of his colleagues have been working on; for the moment, there are still no answers to questions about generalization, a theory of teaching, or, even, future directions for research. The authors have not even attempted a "box-score" of the type that Barak Rosenshine (Rosenhine, B., *Teaching behaviors and student achievement*. National

Foundation for Educational Research in England and Wales, 1971) has made famous in the world of research in K-12 settings. Although “bottom-lines” are not discernible, the book undoubtedly serves its own constituents and is of potential interest to researchers and teachers in post-secondary institutions elsewhere.

A. MacKay
Centre for Research in Teaching
The University of Alberta

GLADLY WOULD HE TEACH. *By John W. Chalmers.* Edmonton, Alberta: Alberta Teachers' Association Educational Trust, 1978, 197 pp., \$5.00.

The story of Milton Ezra LaZerte is almost synonymous with the entire history of education in Alberta. Described as an “educational dynamo,” he served as a public school teacher, principal, school inspector, teachers' association president, university instructor, dean of education, trustee, chairman, and alderman. He was a member of most local, provincial and national educational bodies and, as the author of this biography says, “tended to become president of every organization he joined.”

LaZerte died on February 2, 1975, at the age of 90. The Education Society of Edmonton marked its Golden Jubilee by commissioning this biography to honour “one of its most distinguished members and one of Canada's most outstanding educators.” John Chalmers, the author, is a University of Alberta education professor emeritus who has served in many fields of education in the province and who has written several books.

LaZerte, with a B.A. in mathematics and physics, began his teaching career in 1909 in Ontario, the following year moving to Calgary, Alberta, where he took his teacher training. (Those were the days when the normal school teacher training course was just four months, with practice teaching limited to an almost irrelevant scant eight lessons in a city classroom.) Until 1917, when he joined the army, he was in turn teacher, principal, and school inspector in various Alberta centres. Following the war, in 1919 he returned to the field of education in Alberta. In 1920, he enrolled as a student at the University of Chicago which was then a hotbed of the new “progressive education” of John Dewey. His studies there, particularly under psychologist Charles Judd, profoundly affected his future, convincing him that the education of students in Alberta was not education at all. He became committed to the idea of professionalization of teachers and looked ahead to the time when all teachers would be university trained.

The dream of a full-fledged school of education at the University of Alberta became a reality in 1929, when it opened with a student body of seven with LaZerte as head. The next few years were an uphill struggle for the school of education, partly because its training term was longer than that of the normal schools (in those Depression days, it was usually essential to get a paying job as quickly as possible). The 1930's, however, brought many educational changes and by 1940, the school of education was transformed into a College of Education with LaZerte as first Dean.

Despite nominal retirement at 65, LaZerte continued his creative professional career for another incredible 25 years as an academic activist. In addition to being a member of many local educational groups, he became a prime mover in Canadian educational associations, the Canada-U.S. Committee of Education, and also was a Canadian delegate to the international UNESCO Conference in 1950.

Throughout his career, he maintained his early interest in mathematics, and continually pressed for educational research. He wrote texts and articles, and invented a “problem cylinder” which predated Cusinaire and other mechanical and audiovisual teaching methods. Inclined to be a bit stiff outwardly, always wearing his black derby hat (even on a duck hunt!), LaZerte nevertheless had a tremendous influence on Canadian education. Chalmers, like most writers of commissioned biographies, has undoubtedly stressed the positive side of his subject’s life. LaZerte was, says Chalmers,

A man for all seasons: peace and war, economic prosperity and depression, educational progress and stasis. But most of all he was a man for the future. . . . He prophesied, warned, exhorted, planned and built for a better tomorrow.

Wilfrid J. Bennett
West Vancouver School Trustee

THE PROCESS OF THINKING. *By Marc Belth.* New York: David McKay, 1977, 234 pp.

Marc Belth, in his “Preface,” lists an ambitious set of questions as the basis of the book’s explorations. Predictably, only three at the most are given detailed consideration. These are: (i) “What are the instruments of the thinking process?” (ii) “What can be said about the thinking process so that it can be learned?” (iii) “What are the impediments to thinking, to learning, to learning how to think?” (p. xi). Belth’s motivation seems to be derived from the belief that even with our best intentions, the world’s problems cannot be solved without adequate thinking, and since the characteristics of adequate thinking can be taught and learned, it is important to attempt to focus on these characteristics with as much clarity as possible. The main burden of the book is then to identify model-making and metaphor and analogy creation as the essential characteristics of the thinking process in a number of fields — poetry, history, and the physical and social sciences. Many common and problematic assumptions are tackled en route. Among them are the belief that improvement of thought relies upon the elimination of metaphor, the belief that we can have direct (metaphor-free) knowing, and the belief that thought, being private, cannot be an object of analytic examination.

Belth, therefore, is not simply describing or explaining the thinking process as such. He is concerned with good thinking. And since he defines thinking as “the process of analogizing” (p. 5), “good” thinking then becomes the creation and employment of analogies, metaphors, and models which are “better” than those previously employed (p. 76). In short, to the usual assumption that improved thought is at least truer thought, Belth wishes to add the detail that truer thought means truer analogies. Crucial, then, must be his account of the way in which we can tell whether one analogy is truer than another. However, this explicitly epistemological discussion, in my opinion, could be more thoroughly developed. The prevailing assumption seems to be that the meaning of truth is correspondence (see e.g., p. 64), though this is not discussed in the context of other possibilities. So far as the justification of the acceptance of a model as true, or truer than another, is concerned, the discussion in the context of history and social science for the most part assumes the obvious correctness of the theory of verification by observation. This assumption has produced some of the hottest and most sophisticated debate in recent philosophy of science. Failure to assess this debate and draw conclusions which avoid the crucial stumbling blocks detracts significantly from the value of the book. Nowhere do I find even the attempt to construct a clear and decisive

argument settling once and for all the question whether producing “better” models amounts to anything more than creating more interesting fantasies about a hypothetical “world.” There is at least one suggestion of a pragmatic criterion (p. 154), but there is no account of why we should interpret truth in pragmatic terms or why, alternatively, we should give up truth and opt for usefulness which we can gauge. I am bothered in this respect by the question: Is it true that a proposed model is more useful? Without such an explicit construction of a theory of verification, the author’s challenge of entities outside the scope of perceptual observation and physicalistic theory (labelling them, for example, “mysterious metaphysical existences” [p. 207]) is questionable.

Requiring further discussion is Belth’s adopted meaning of “thinking.” We are told that “the only time that models are not in use are those occasions when no thinking occurs” (p. 60). If “thinking” has been defined as model-making (as it has), this must be circular and quite uninformative, that is, unless there is genuinely and obviously no reason to term as “thinking” anything which does not involve creating and (critically) using models. But there is, I think, an excellent reason to widen the scope of the definition of thinking. For where the emphasis is on good thinking and on improving thinking, there is an appeal to the crucial mental ability of “perfecting” a model by using it “reflectively” and “critically.” This ability, however, must surely be beyond the use of models, for the mere substitution of one model for another could not by itself constitute a better description or explanation. The conclusion would seem to be that thinking improves on account of that aspect of its activity which involves direct (model-free) insight into the similarity of form between model and events investigated with the model. This very ability which Belth would seem to require in his theory to render intelligible his concept of improvement of thought, he explicitly denies. He inconsistently assumes such insight in the concept of model-perfecting, and when, for example, he uses the concept of “raw” data (p. 198); his stated position, however, is that we cannot “look at nature directly” (p. 61). And because model-free insight enters the discussion only tacitly, a very important topic in any consideration of thinking is ignored — the great and genuine mystery of how, as subjects, we can have an objective grasp on the character of the world. The only objectivity he offers without adequate argument is that of the public examinability of models, and that is unsatisfactory since the expression of any cultural belief of universal scope may be publicly examined and an indefinite amount of evidential support claimed, as Polanyi and Feyerabend, for example, have shown (Polanyi, 1962, ch. 9; Feyerabend, 1961).

Bearing in mind the unacceptable narrowing of the meaning of ‘thinking,’ and the actual scale of achievement in the book, I would think that a more accurate title would be: “The Role of Model, Metaphor and Analogy in Thinking.” That models are indispensable vehicles in the development of thought is a thesis amply defended by Belth, but no satisfactory case is made for the identification of thought with the process of model invention and use. Without this, the title as it stands is misleading.

If, therefore, the concept of model-improvement involving truth value is to stand and if that logically requires some form of direct insight into the nature of things, two interesting problems arise. Firstly, if we can know the world directly, why do we need models, metaphors, and analogies? Secondly, if we have no direct knowledge, how can we check the truth value of our models? It could, I think, have been extremely illuminating about the process of thought if Belth had slightly, but significantly, reoriented his wealth of material to address these questions.

Belth makes it clear right at the start that his concern is to encourage the

attempt to understand thought in such a way as to liberate its fullest possible effectiveness in the actual world of present existence (pp. xviii-xxi). With the soundness of this intent it would be hard to disagree, but there is some doubt about the wisdom of Belth's own approach to the solution of this problem of relevance and of concreteness. I suspect that for the average graduate or upper undergraduate student, there would be adequate concrete illustration only in the chapter on poetic thought. For most, the mass of abstract generalizations could easily be lacking significant meaning. Even in the chapter on scientific thinking, the attempt to work concretely in terms of the example of Melzack's theory of the perception of pain is rendered largely ineffective, for at the very point where the effectiveness of the example is to be realized — the explicit connection of the details of the example with the list of abstract elements of a scientific model — we are merely told that "each of these elements can be recognized in the illustration" (p. 194). But can they? I very much doubt it, except by readers whose sophistication in the understanding of contemporary model theory is close to Belth's own.

There is awareness in the presentation of ideas that if interest in understanding thinking is to occur, the vagueness, paradoxicality, and apparent groundlessness of some of our most pervasive assumptions about the process must be exposed so sharply that we feel we cannot proceed without a resolution of the problems. The supreme educational art of the provocateur is required here, of which Socrates' dialectic is a fine example. For the most part, however, Belth drops his controversial theses far too soon, settling for what is almost a mere listing of alternatives before moving on. The effect is more that of an amassing of information. The one sustained and exhaustively illustrated argument is that all thinking involves models, metaphors, and analogies, but the lack of a clear sense of contrast reduces its effectiveness. For we are never shown unmistakably enough through quotational illustration that anyone seriously doubts the truth of this thesis nor what, through a relentless drawing of unavoidable implications, the unwelcome consequences of believing the reverse might be. Much is suggested, hinted at, and implied in this respect, but this is not enough to accomplish the special engagement of interest at which Belth aims.

As previously argued, the main theme itself — concerning the impossibility of thought without model, metaphor, and analogy — though exhaustively illustrated, is inadequately demonstrated on crucial philosophical points. Too few of the questions raised receive adequate treatment. Unfortunately, the introduction of too many questions hinders the author's fine intention and reduces the effectiveness of his obviously impressive background scholarship.

Foster N. Walker
Department of Educational Foundations
The University of Alberta

References

- Feyerabend, P. K. *Knowledge without foundations*. Ohio: Oberlin College, 1961 (mimeographed).
Polanyi, M. *Personal knowledge*. Chicago: University of Chicago Press, 1962.

LIST OF BOOKS RECEIVED

- Bassett, G. W., Watts, B. H., & Nurcombe, B. *Individual differences*. Winchester, Mass.: Allen & Unwin Inc., 1978, 284 pp., \$23.75 (cloth), \$12.95 (paper).
- Beard, R. M., Bligh, D. A., & Harding, A. G. *Research into teaching methods in higher education*. Surrey, England: Society for Research into Higher Education, 1978, 150 pp.
- Beck, C., et al. *Final report 1967-77: The moral education project (Year 5)*. Toronto: Ontario Institute for Studies in Education, 1979, \$3.00.
- Belth, M. *The process of thinking*. New York: David McKay, 1977, 234 pp.
- Berkeley, H., Gaffield, C., & West, W. G. (Eds.) *Children's rights: Legal and educational issues*, Toronto: Ontario Institute for Studies in Education, 1978, 177 pp.
- Billing, D. *Course design and student learning: Papers presented at the Society's thirteenth annual conference, 1977*. Surrey, England: Society for Research into Higher Education, 1978, 167 pp.
- Blaug, M. *Economics of education: A selected annotated bibliography (3rd ed.)*. Elmsford, N.Y.: Pergamon Press Ltd., 1978, 428 pp.
- Brown, E. K. *Foundations of educational evaluation*. Ardmore, PA: Dorrance & Company, Inc., 1978, 192 pp., \$10.00.
- Canadian Education Association. *Educational research and policy formation*. CEA/CERA Research Conference, 1977. Toronto, 1978, 128 pp., \$5.00.
- Canadian Education Association. *Requirements for teaching certificates in Canada*. Toronto, 1978, \$2.50.
- Carson, R. B., & Friesen, J. W. *Teacher participation: A second look*. Washington: University Press of America, 1978, 103 pp., \$6.75.
- Chalmers, J. W. *Gladly would he teach*. Edmonton: Alberta Teachers' Association Educational Trust, 1978, 197 pp., \$5.00.
- Clough, D. B., & Clough, B. M. *A handbook of effective techniques for teacher aides*. Springfield, Ill.: Charles C. Thomas, 1978, 189 pp., \$10.75.
- Collier, G. (Ed.) *Evaluating the new B.Ed.* Surrey, England: Society for Research into Higher Education, 1978, 120 pp.
- Cremin, L. A. *Traditions of American education*. Don Mills, Ontario: Fitzhenry & Whiteside Limited, 1977, 172 pp., \$5.50.
- Downey, M., & Kelly, A. V. *Moral education: Theory and practice*. New York: Harper & Row, 1978, 226 pp., \$7.50.
- Fowler, W. *Day care and its effects on early development*. Toronto: Ontario Institute for Studies in Education, 1978, 107 pp., \$7.75.
- Fowler, W. *Guides to early day care and teaching* (Supplement to *Day care and its effects on early development*). Toronto: Ontario Institute for Studies in Education, 1978, 83 pp., \$3.50.
- Gagliardi, R. *The mathematics of the energy crisis*. Westmont, NJ: Intergalactic Publishing Company, 1978, 72 pp.
- Gayfer, M. *An overview of Canadian education: Second edition*. Toronto: Canadian Education Association, 1978, 43 pp., \$2.50.
- Gayfer, M. *Canadian approaches to school health education and services*. Report of a CEA survey. Toronto: Canadian Education Association, 1978, 64 pp., \$3.00.

- Glenn, J. A. (Ed.) *The third R: Towards a numerate society*. Don Mills, Ontario: Fitzhenry & Whiteside Limited, 1978, 140 pp., \$6.95.
- Hassam, A. S. *The learning process: How to pass examinations*. Hicksville, N.Y.: Exposition Press, Inc., 1978, 79 pp., \$5.50.
- Hétu, J.-C. *Stratégie d'enseignement des nombres entiers naturels*. Montréal: Les Presses de l'Université de Montréal, 1978, 150 pp., \$12.75.
- Hillerich, R. L. *A writing vocabulary of elementary children*. Springfield, Ill.: Charles C. Thomas, 1978, 295 pp., \$13.75.
- Hodgetts, A. B., & Gallagher, P. *Teaching Canada for the '80s*. Toronto: Ontario Institute for Studies in Education, 1978, 136 pp., \$5.95.
- Kelly, A. V. *The curriculum: Theory and practice*. New York: Harper & Row, 1977, 202 pp., \$6.35.
- Kurtz, A. K., & Mayo, S. T. *Statistical methods in education and psychology*. New York: Springer-Verlag New York Inc., 1978, 540 pp., \$12.00.
- Lewy, A. (Ed.) *Handbook of curriculum evaluation*. Toronto: Ontario Institute for Studies in Education, 1977.
- Mussen, P. H., Conger, J. J., & Kagan, J. *Child development and personality. Fifth edition*. New York: Harper & Row, 1979, 577 pp., \$15.95.
- Nyberg, V. R., & Lee, B. *Evaluating academic achievement in the last three years of secondary school in Canada*. Toronto: Canadian Education Association, 1978, 60 pp., \$3.00.
- Prokos, G. *This is about attendance counselling*. Toronto: Canadian Education Association, 1978, 23 pp., \$1.50.
- Ravitch, D. *The revisionists revised*. New York: Basic Books, 1978, 194 pp., \$11.65.
- Richards, J. *Classroom language: What sort?* Winchester, Mass.: Allen & Unwin Inc., 1978, 154 pp., \$16.25 (cloth), \$7.50 (paper).
- Ricker, E. (Ed.) *Education and development in Atlantic Canada*. Halifax, N.S.: Dalhousie University, 1978, 381 pp., \$8.00.
- Rosier, M. J. *Early school leavers in Australia*. Stockholm, Sweden: Almqvist & Wiksell International, 1978, 198 pp.
- Sawyer, D. C. (Ed.) *Directory of education studies in Canada, 1976-1977*. Toronto: Canadian Education Association, 1978, 314 pp., \$6.00.
- Swain, M., & Harley, B. (Eds.) *Working papers on bilingualism, Issue No. 16*. Toronto: Ontario Institute for Studies in Education, 1978, 119 pp.
- Swain, M., & Harley, B. (Eds.) *Working papers on bilingualism, Issue No. 17*. Toronto: Ontario Institute for Studies in Education, 1979, 198 pp.
- Weizman, R., Brown, R., Levinson, P. J., & Taylor, P. A. (Eds.) *Piagetian theory and the helping professions*. Los Angeles: University of Southern California, 1978, 432 pp., \$12.00.

PUBLICATIONS

FACULTY OF EDUCATION,
UNIVERSITY OF ALBERTA
EDMONTON T6G 2G5

The Alberta Journal of Educational Research

AJER is a quarterly journal devoted to the dissemination, criticism, interpretation and encouragement of all forms of systematic enquiry into education and fields related to or associated with education. Published in March, June, September and December; subscription is \$8 per year. Address communications to the Editor, *AJER*, 732 Education South.

Elements

Elements endeavours to help educators translate theory and findings from current research into classroom practice. Each issue focuses on a specific subject area, topic or theme. Published monthly September through April; subscription is \$4 per year. Address communications to the Editor, *Elements*, Department of Elementary Education, 539 Education South.

Indian-Ed

Indian-Ed publishes articles, presents abstracts of research and reports news items related to the education of native peoples. Published four times per year; subscription is \$3. Address communications to the Editor, *Indian-Ed*, Department of Educational Foundations, 5-109 Education North.

The Canadian Administrator

The **CA** serves administrators at all levels of educational organizations. Articles selected for publication are intended to familiarize practitioners with the findings and implications of research, and with analyses of current policies and practices. Published eight times per year October through May; subscription is \$5 per year. Address communications to Editor, *CA*, Department of Educational Administration, 7-104 Education North.

The Canadian Journal of Sociology Cahiers canadiens de sociologie

Editors

Richard Ericson
Andrew Harrell
Lyle Larson
Nico Stehr

Editorial Board

Gerald Fortin
Edward Gross
Stanley Lieberman
Mildred Schwartz
Dennis Wrong

Associate Editors

T.R. Balakrishnan—University of Western Ontario
Wallace Clement—McMaster University
James Curtis—University of Waterloo
Leo Driedger—University of Manitoba
Larry Felt—Memorial University
Dennis Forcese—Carleton University
Martha Foschi—University of British Columbia
Herbert Gamberg—Dalhousie University
Timothy Hartnagel—University of Alberta
R. Alan Hedley—University of Victoria
Harry H. Hiller—University of Calgary
John D. Jackson—Concordia University
J. Paul Johnston—University of Alberta
Ronald D. Lambert—University of Waterloo
Eugen Lupri—University of Calgary
William Michelson—University of Toronto
James C. Moore—York University
James Richardson—University of New Brunswick
Brigitte Schroeder-Gudehus—Université de Montréal
Dorothy Smith—Ontario Institute for Studies in Education
Malcolm Spector—McGill University
Lorne Tepperman—University of Toronto
Roy Turner—University of British Columbia
Kenneth Westhues—University of Waterloo

The Canadian Journal of Sociology is published at
The University of Alberta. Individual subscriptions are
\$18 per year. Subscriptions from institutions and libraries are
\$36 per year. Requests concerning manuscripts,
advertising and subscriptions should be sent to
The Canadian Journal of Sociology, Department of Sociology,
The University of Alberta, Edmonton, Alberta, Canada T6G 2H4.

Some recent and
forthcoming articles from the
quarterly CJS: **Bernd Baldus**—
Social control in capitalist
societies; **Alan Blum and Peter McHugh**—
A Sense of Place; **Merlin Brinkerhoff**—
Women who want to work in a man's world;
S.D. Clark—Sociology in Canada; **Harold Fallding**—
Mainline protestantism in Canada and the United
States; **Harriet Friedmann and Jack Wayne**—Depend-
ency theory; **Richard Henshel and Robert Silverman**—
Perception and criminal process; **I.C. Jarvie**—Nationalism
and the social sciences; **Niklas Luhmann**—Differentiation of
society; **Neil Mackinnon and Gene Summers**—Homogeneity and role
consensus; **Wolfgang Mommsen**—Max Weber as critic
of Marxism; **John F. Myles and Aage Sorensen**—Elite and status attain-
ment models; **John Rex**—Value relevance, scientific laws, and ideal types;
James Richardson—Education and social mobility; **Paul Stevenson**—Frustration,
structural blame and leftwing radicalism; **Lorne Tepperman**—Deviance as a search process.

PREPARATION OF MANUSCRIPTS

1. All manuscripts must be typewritten, double spaced, and submitted in duplicate. An abstract of approximately 100 words in length, typed on a separate page, should be provided.
2. Tables must be numbered in Arabic numerals with the word 'Table' centered and in capital letters, e.g., TABLE 4. The heading of the table is to be centered below and typed in capitals. The format of tables should conform to the specifications in the APA Publications Manual.
3. Graphs and charts should be used only if essential. They must be carefully prepared on separate sheets in India ink, ready for reproduction. Graphs must be properly labelled using Arabic numerals, e.g., Figure 3.
4. Each table or figure should be presented on a separate page. The position of tables and graphs should be clearly indicated within the text by inserting at the relevant point the phrase (Insert Table 2 here).
5. References should appear in parentheses following the reference citing the author's name (unless the name appears in the text), the year of publication, and page number if appropriate. For direct quotations, the reference should be cited and the page number given in brackets before the final punctuation of the quotation. The references should be listed alphabetically by author's last names at the end of the manuscript under the heading, *References*.
6. Explanatory notes, numbered consecutively and identified in the text with a superscript, may be included under the heading of *Notes*. They should be double spaced and placed at the end of the manuscript immediately preceding the *References*. The citing of references and quotations in the *Notes* should conform to the procedures outlined in No. 5 above.
7. Spelling shall conform to the *Oxford English Dictionary*, except in the case of direct quotations which must conform to the original. Editorial alterations will be made if necessary.
8. In matters of style, the APA Publications Manual is considered definitive.

